# A METHODOLOGY FOR QUANTIFYING AND IMPROVING PAVEMENT CONDITION ESTIMATION AND FORECASTING BY INTEGRATING SMARTPHONE AND 3D LASER DATA

A Dissertation
Presented to
The Academic Faculty

By

Anirban Chatterjee

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Computational Science and Engineering
School of Civil and Environmental Engineering

Georgia Institute of Technology

August 2019

**A METHODOLOGY FOR QUANTIFYING AND IMPROVING PAVEMENT CONDITION ESTIMATION AND FORECASTING BY INTEGRATING SMARTPHONE AND 3D LASER DATA**

Approved by:

Dr. Yi-Chang (James) Tsai, Advisor
School of Civil and Environmental Engineering
*Georgia Institute of Technology*

Dr. Adjo Amekudzi-Kennedy
School of Civil and Environmental Engineering
*Georgia Institute of Technology*

Dr. Zhaohua Wang
School of Architecture
*Georgia Institute of Technology*

Dr. Ümit V. Çatalyürek
School of Computational Science and Engineering
*Georgia Institute of Technology*

Dr. Richard Vuduc
School of Computational Science and Engineering
*Georgia Institute of Technology*

Date Approved: July 12, 2019

*For nush.*

## ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisor, Dr. Yi-Chang Tsai, for his support. He gave me the opportunity to tie my background in transportation engineering to my passion for computer science for which I shall forever be grateful. His guidance and confidence in me has helped to make this research possible.

I would also like to thank my committee members, Dr. Adjo Amekudzi-Kennedy, Dr. Ümit V. Çatalyürek, Dr. Richard Vuduc and Dr. Zhaohua Wang for serving as my committee members and providing valuable feedback which has helped to shape this dissertation.

I would also like to express my appreciation to our research group, especially Dr. Zhaohua Wang and Yi-Ching Wu for their constant support and feedback.

Finally, I am grateful to my family, for their support and confidence in me.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## SUMMARY

The US federal and state governments spent USD 72.7 Billion in 2014 on the operation and maintenance of highway infrastructure. Even then, there exists a lack of funding for the adequate maintenance of these highways. To optimize the use of resources for infrastructure maintenance, regular road infrastructure condition surveys are required. Automated road condition surveys involve the use of survey vehicles to collect road infrastructure condition data and distress detection algorithms to automatically assess the infrastructure condition. However, existing approaches for automated road condition surveys are subject to several shortcomings.

3D laser technology has become the mainstream technology for pavement condition assessment, used by over 21 State Departments of Transportation (DOTs) in the US. However, its high cost makes frequent (quarterly or monthly) inspection impractical. Low-cost vehicle mounted sensors, such as smartphones can crowdsource pavement condition data much more frequently. However, the range of distresses covered in existing research on low-cost vehicle-mounted sensor based pavement condition estimation is limited, the estimates are highly sensitive to user context, and there is no confidence level associated with the pavement condition estimates making it difficult to separate good and bad estimates.

This study aims to combine data from accurate but expensive 3D laser scanners and inexpensive and frequently used smartphones for improving pavement condition estimation and forecasting using both technologies. This study presents:

1. A methodology for registering both 3D laser data and smartphone sensor data onto a common GIS model of the road network;

2. A methodology for single-run pavement condition estimation using smartphone data trained using labeled data from a 3D laser scanner;

3. A methodology for combining multiple-run pavement condition estimates from both

smartphones and 3D laser scanners with an associated confidence level; and

4. A methodology to improve pavement condition forecasting using 3D laser data by combining updated evidence obtained from crowdsourced smartphone data.

Data was collected on a test route consisting of diverse pavement conditions and road classifications to test and validate the proposed methodologies. The data registration methodology was validated by observing a large peak in cross-correlation between pavement distress metrics calculated from multiple runs registered using the proposed methodology. The methodologies for pavement condition estimation and forecasting were validated by observing low median error for International Roughness Index (IRI) estimation and forecasting. The proposed method for associating a confidence level with each estimate was able to successfully separate low and high error estimates derived from smartphone data, allowing transportation agencies to use smartphone derived pavement condition estimates with low error.

This research will strongly improve the utility of both 3D laser technologies and smartphones for pavement condition estimation and forecasting.

# CHAPTER 1
# INTRODUCTION

## 1.1 Background

Regular road infrastructure maintenance is vital for improving road safety while decreasing user costs. There are 4,184,471 centerline miles of public roads in the US (FHWA, 2017a). The US federal and state governments spent USD 82.72 Billion in 2016 for the operation and maintenance of highway infrastructure (Musick, 2018). Even then, the ASCE estimates a backlog of USD 420 Billion in capital needs in the US for repairing existing highways (ASCE, 2017). In 2018, the Trump administration released an infrastructure proposal calling for investments in infrastructure of USD 1.5 trillion, with the federal government directly spending USD 200 Billion over the next 10 years (The White House, 2018).

Infrastructure asset management is vital for optimizing the use of limited resources for the maintenance of road infrastructure, and the essential first step for infrastructure asset management is to accurately and continuously assess the condition of the road infrastructure. Detailed, regular road condition assessment is also a federal requirement under the FAST ACT. Thus, an accurate and cost-effective method for frequently assessing the road infrastructure is urgently needed.

Manual road infrastructure condition assessments are still used, in which engineers manually assess the road infrastructure condition through on-foot or vehicle surveys. These surveys are typically infrequent (once a year) and cover a small portion of the total centerline miles of roadway. For example, the current GDOT approach to crack detection is to manually measure the length of cracking on a 100 ft. representative section of each mile of roadway (GDOT, 2007). This approach is time-consuming, expensive and error-prone (Rada, Bhandari, Elkins, and Bellinger, 1997). Additionally, using a 100 ft. section to

represent a 1-mile road segment means a sampling of less than 2% of the centerline miles of roadway is used to represent the condition of the entire road network.

Recent advancements in automatic road infrastructure condition assessment using specialized vehicles outfitted with sensors allow the road infrastructure condition to be automatically determined at highway speeds (Tsai and Wang, 2013). 3D laser technology has become the mainstream technology for state DOTs to collect a high-resolution 3D image of the pavement surface with full lane coverage, which is then used for automatic extraction of pavement distresses, including cracking, rutting, raveling, potholes and IRI. This 3D laser technology is used by more than 21 state DOTs to collect road infrastructure data. The TPF-5(299) pooled fund study, managed by the Federal Highway Administration (FHWA), has been established nationally to improve data collection and analysis using this technology (FHWA, 2017b). However, the sensors and vehicles for automated road condition assessment have a high cost. Therefore, the data is often collected yearly or once every two years. With such an infrequent data collection interval, it is difficult to predict the pavement condition accurately in between surveys, especially when a pavement enters the end of its life with severe deterioration, during which the pavement condition needs to be identified frequently to apply the appropriate treatment at the right time to ensure the roadway safety. Thus, a method to improve pavement condition forecasting between infrequent 3D pavement data collection is needed.

The feasibility of low-cost vehicle mounted sensors and smartphones for road condition assessment, especially for estimating the International Roughness Index (IRI), has also been explored (Dawkins, Bishop, Powell, and Bevly, 2011; Sauerwein and Smith, 2011; Flintsch, Valeri, Katicha, Leon Izeppi, and Medina-Flintsch, 2012; Zang, Shen, Huang, Wan, and Shi, 2018; Eriksson, Girod, Hull, Newton, Madden, and Balakrishnan, 2008; Chen, Lu, Tan, and Wu, 2013). IRI is a measure of the roughness of the pavement surface (Sayers and Karamihas, 1996), which affects ride quality. Smartphone based pavement condition assessment methods have the advantages of low cost and frequent pavement

condition assessment over the 3D laser technology. These low-cost methods enable the crowdsourcing of road condition information, which make these methods easily scalable both temporally and spatially. Engineers are interested in using automated road condition assessment methods to augment manual road surveys by helping to shortlist road segments to be manually surveyed (Zeng, 2014). However, these low-cost methods have several disadvantages. First, the existing methods focus on only two pavement distresses: pavement roughness and potholes. Second, user context can affect the signal response from the smartphone for the same road condition (Zeng, Park, Smith, and Parkany, 2018). The signal response of smartphones are dependent on more factors than the road condition, such as vehicle model, vehicle speed and position and orientation of the sensor inside the vehicle, which makes it difficult to use them for road condition estimation. Most importantly, existing methods of road condition evaluation using smartphones do not associate a confidence level with the pavement condition estimates. Thus, a mix of good and poor pavement condition estimates are combined in the existing methods, which significantly hinders the adoption of current smartphone based pavement condition assessment methods. Without a confidence level, all estimates have to be considered equally accurate, and have to be either universally accepted or rejected. Thus, smartphones offer an effective solution for frequent and cost-effective road condition assessment as compared to the previously listed approaches. Unfortunately, the 3 major disadvantages listed above decreases their viability as a useful solution. Thus, an improved method for pavement condition assessment using smartphone data which mitigates the current shortcomings is urgently needed.

Our research team at Georgia Tech has developed the Georgia Tech Sensing Vehicle (GTSV), a USD 1 Million investment, which is integrated with a 2D imaging system, 3D laser scanner system (LCMS, manufactured by INO/Pavemetrics), mobile light detection and ranging (LiDAR), high-resolution Inertial Measurement Unit (IMU), high-accuracy differential GPS, and a high-frequency Distance Measuring Instrument (DMI) (shown in Figure 1) through two national demonstration projects (Remote Sensing and GIS-enabled

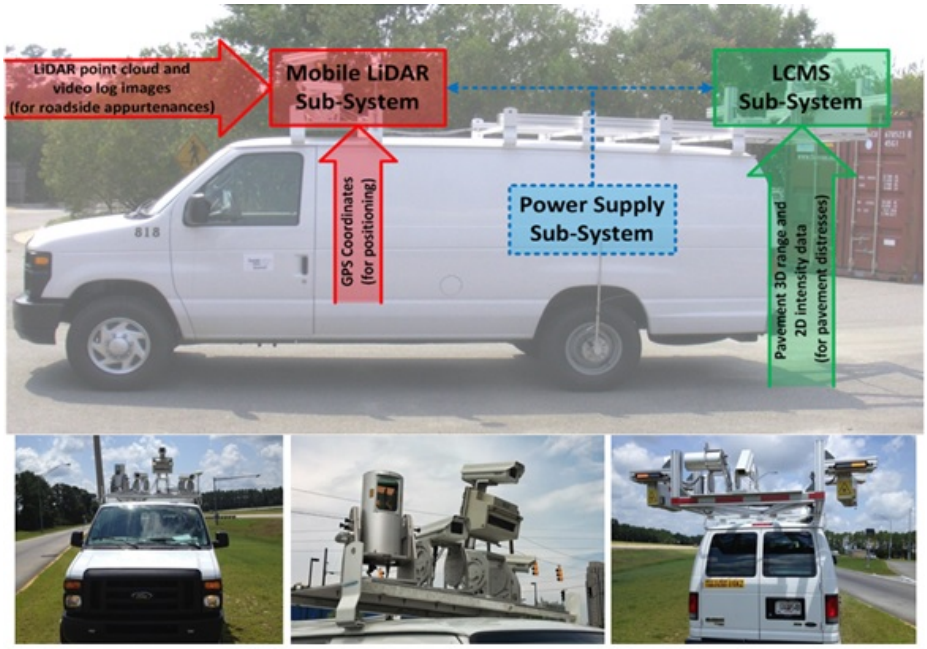



Figure 1.1: Georgia Tech Sensing Vehicle (Tsai and Wang, 2014)

Asset Management System, RS-GAMS Phases I and II) sponsored by the USDOT Office of the Assistant Secretary for Research and Technology (USDOT/OST-R) since 2010 (Tsai and Wang, 2013; Tsai and Wang, 2014).

The GTSV has been intensively applied by Georgia Tech researchers to conduct research and for practical applications, including crack detection (Chatterjee, 2017; Tsai and Li, 2012; Jiang and Tsai, 2016), rutting detection (Tsai, Wang, and Li, 2015; Tsai, Li, and Wu, 2013), raveling detection and classification (Tsai and Wang, 2015), pothole detection (Tsai and Chatterjee, 2017), concrete joint faulting measurement (Tsai, Wu, and Ai, 2011) and project-level micro-milling pavement surface texture construction quality control (Tsai, Wu, and Lewis, 2013). It has been used to support my MS thesis on automatic pavement crack detection using 3D technology (Chatterjee, 2017) and has been continuously used to support my PhD study.

Our research group has the unique position of having collected and processed extensive road infrastructure data using the Georgia Tech Sensing Vehicle equipped with a 3D laser scanner, as well as estimating road condition using low-cost methods such as smartphones.

This brings a research opportunity to combine signals from both smartphones sensors and laser scanners to create a hybrid road condition assessment system which has the advantages of both sensors:

- The laser scanner data helps to provide a much more granular pavement condition information with multiple types of distresses which can be used as labels to train models which use the smartphone sensor data as input to estimate the pavement condition more accurately with an associated confidence level, even at locations or timestamps where laser scanner information is unavailable.

- The smartphones then offer a cost-effective, scalable method for frequent pavement condition estimation, which can be used to enhance the forecasting of the pavement condition between infrequently collected 3D pavement data.

Thus, this hybrid approach has the potential to solve the shortcomings of both laser scanner based methods and smartphone based approaches.

## 1.2 Research Objective

The objective of this study is to develop a methodology to improve pavement condition estimation and forecasting using crowdsourced low-cost, frequent smartphone sensor data and infrequent high-resolution laser scanner data. The outcomes of the proposed study will improve the accuracy of smartphone-based pavement condition evaluation for local transportation agencies using low-cost methods, and improve 3D pavement data based pavement condition forecasting for larger transportation agencies using frequently crowdsourced smartphone data. The following research objectives will be completed to fulfill the stated objective:

- To develop a methodology for quantifying and improving pavement condition estimation using smartphone data and sparse 3D pavement data. The proposed methodology will associate a confidence level with the pavement condition estimations.

- To develop a methodology for improving 3D pavement data based pavement condition forecasting by taking advantage of low-cost, frequently collected smartphone data.

- To conduct comprehensive case studies for the validation of the developed methods.

The following are the methods to be developed in the proposed methodologies:

- A method for registering multiple series of smartphone sensor data and laser scanner data to a GIS model of a road network will be developed. The method will register collected data both spatially to points on the road network and temporally to corresponding data from other series.

- A deep learning based method for estimating the road condition using smartphone sensor data, which addresses the shortcomings of existing methods of road condition estimation using smartphones, will be developed. The model of the proposed method will be trained with road condition information determined using a laser scanner.

- A method for combining multi-run pavement condition estimations from both smartphone and 3D pavement data will be developed. The combined pavement condition estimations will have a confidence level associated with them.

- A method to improve pavement condition forecasting using crowdsourced smartphone will be proposed. The pavement condition forecasts will include an associated level of confidence to improve its utility.

## 1.3 Thesis Organization

Chapter 2 reviews existing literature on pavement condition assessment using both high-cost specialized sensors including 3D laser scanners and low-cost sensors such as smartphones and their current advantages and shortcomings. Chapter 3 explains the proposed

methodology and validation for data collection and registration of 3D pavement data and smartphone data. Chapter 4 describes the proposed methodology and validation of single-run pavement condition estimation using smartphone sensor data. Chapter 5 explains the proposed methodology and validation for multiple-run pavement condition estimation and pavement condition forecasting with an associated confidence level. Chapter 6 consists of case studies performed to evaluate the performance of the proposed methodologies in a real-world setting. Finally, chapter 7 concludes the study, identifies its contributions and recommends future research work.

# CHAPTER 2
# LITERATURE REVIEW

Almost every transportation agency in US and Canada use some automated means of road infrastructure condition assessment (McGhee, 2004), especially for measuring pavement roughness and rutting. This literature review first summarizes existing automated high-cost specialized sensors for measuring the structural and functional performance of pavements. This is followed by a review of low-cost vehicle-mounted sensor based methods of road infrastructure condition assessment, including smartphones. Finally, the research needs are summarized.

Pavement condition metrics can describe the structural performance or the functional performance of the pavement. Structural performance metrics measure the capacity of the pavement to perform its intended function before structural failure. Functional performance metrics measure the pavements performance from the perspective of the road users, by emphasizing on ride quality and user comfort.

## 2.1 Structural Performance Measurement Methods

Falling weight deflectometers (FWDs) (figure 2.1) work on the principle of applying an impulse load on the pavement and measuring the strain response of the pavement surface at fixed distances from the load to estimate the pavement condition. The road condition is quantified using structural performance indicators such as layer moduli, deflection basin parameters, strain responses (Xu, Ranji Ranjithan, and Richard Kim, 2002) and empirical indices (Elbagalati, Elseifi, Gaspard, and Zhang, 2018; Abd El-Raof, Abd El-Hakim, El-Badawy, and Afify, 2018).

Ground penetrating radars (GPRs) have also been used as a non-destructive method for evaluating road conditions. As the name implies, GPRs measure radio waves reflected off

Figure 2.1: Falling Weight Deflectometer

interfaces between pavement layers to assess their condition. Lai et al. discuss over 30 years of research in assessing civil infrastructure condition using GPRs (Lai, Dérobert, and Annan, 2018). Generally higher frequency waves have higher resolution, at the cost of range. 100-1000 MHz waves as used for analysis of pavements. Air-launched GPRs can be operated without contact with the pavement, enabling highway speed data collection, at the cost of resolution.

For network-level analysis, the road network is generally divided into homogenous sections and FWD/GPR test results obtained within a section are applied to the entire section. Sensors like FWDs and GPRs require a lane of the road to be closed off while the testing is carried out. Thus, they are much more time-consuming than other methods which can assess the road condition continuously at highway speeds.

Traffic speed deflectometers (TSDs) provide a method to apply the same principle of FWDs continuously at highway speeds to analyze road condition (Nasimifar, Thyagarajan, Siddharthan, and Sivaneswaran, 2016; Zofka and Sudyka, 2015). TSDs consist of a trailer

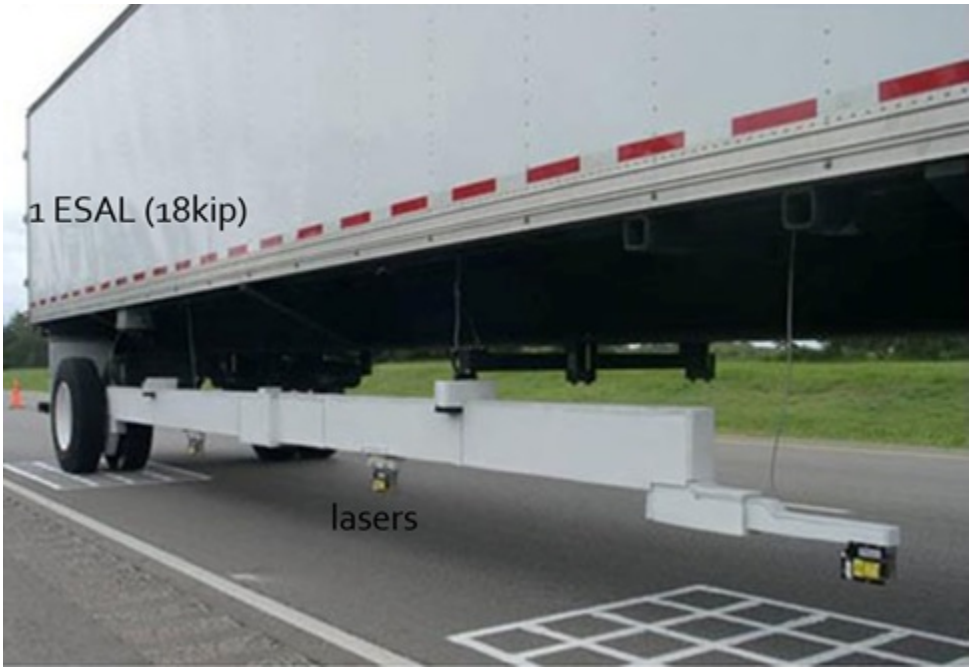



Figure 2.2: Traffic Speed Deflectometer

truck with sensors placed along the length of the trailer. The rear axle of the trailer is calibrated to simulate a standard axle load (18kip SADW). The sensors then continuously measure the response of the pavement due to the axle load as the vehicle travels along the roadway at highway speeds.

It is difficult to develop low-cost methods for structural performance measurement as specialized sensors will be required to measure the response of the pavement to an applied load or pulse wave. Structural performance may not correctly describe the road users experience, which is better described by the presence and severity of surface distresses such as cracking, rutting, raveling, potholes and the roughness of roads. The performance of the road infrastructure from the perspective of the user is referred to as its functional performance, which is commonly used in the US. Methods for measuring the functional performance are described in section 2.2.

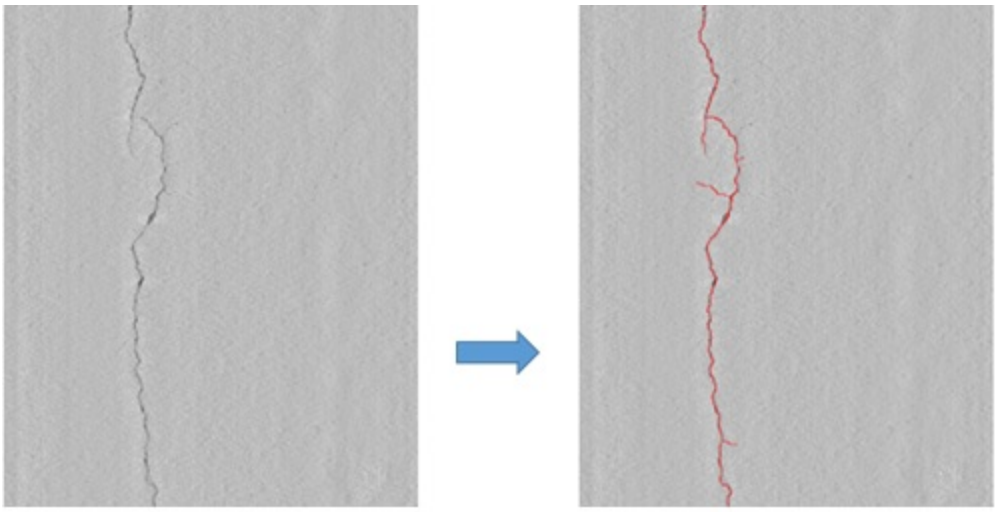


Figure 2.3: Automated crack detection on 3D pavement image data (Chatterjee, 2017)

## 2.2 Functional Performance Measurement Methods

Image based automated road condition survey systems, which can be mounted on vehicles, have been developed since the early 1990s (Wang, 2000), including 3D imaging technologies (Tsai and Li, 2012; Wang, 2004). After capturing image data of the pavement surface, computer vision concepts are applied to extract pavement distresses from the image data, which are used to quantitatively measure the road condition. Computer vision algorithms to extract cracks (Chatterjee, 2017; Tsai and Li, 2012; Jiang and Tsai, 2016; Zhang, Wang, Li, Yang, Dai, Peng, Fei, Liu, Li, and Chen, 2017), rutting (Tsai, Wang, and Li, 2015; Tsai, Li, and Wu, 2013), potholes (Tsai and Chatterjee, 2017; Koch and Brilakis, 2011) and other distresses (Tsai, Wu, and Ai, 2011) from pavement image data is a heavily researched topic. Figure 2.3 demonstrates the detection of pavement cracking from 3D pavement image data.

3D laser scanners have become a mainstream technology for pavement data collection and processing. Laser scanners collect the intensity of emitted light reflected off the pavement surface as well as the depth of the pavement surface from a fixed height (Tsai and Wang, 2013). Figure 2.4 demonstrates the intensity and range images captured by laser scanners. In the range image, the brightness of the pixels is inversely proportional to the depth of the pavement surface at that point from a fixed datum on the vehicle. Thus, darker

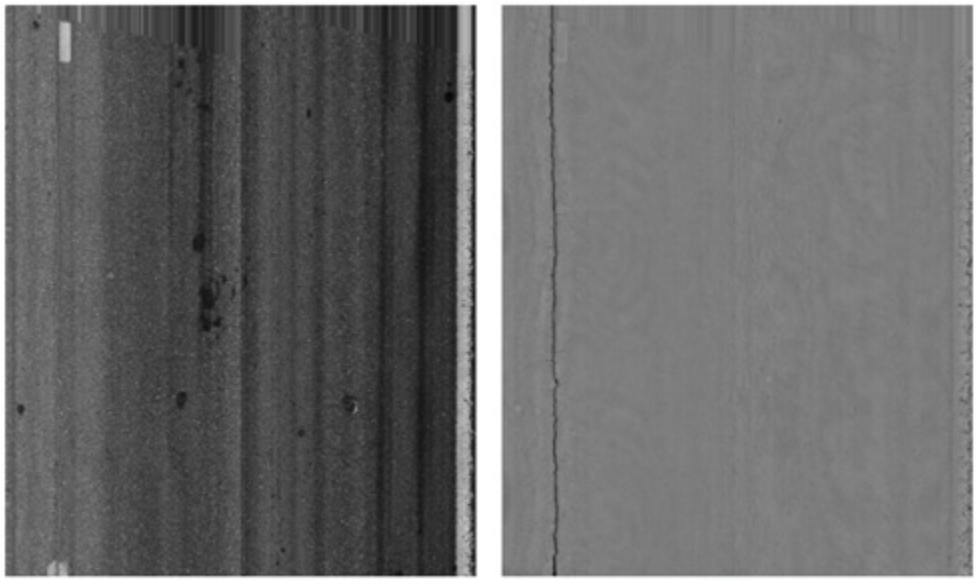



Figure 2.4: Pavement image using a laser scanner: (a) intensity image; (b) range image (Tsai and Li, 2012)

pixels indicate deeper parts of the pavement surface while brighter pixels indicate elevated portions of the pavement surface.

This 3D laser system has demonstrated its advantages over the traditional 2D intensity imaging system (Tsai and Li, 2012). First, the 3D laser-based system is not sensitive to lighting effects (figure 2.5) as it measures the range (i.e. elevation) like other laser and Light Detection and Ranging (LiDAR) devices. Second, some noises, like oil stains and poor intensity contrast, will not interfere with the pavement distress detection algorithms using the acquired range data. Increased attention has been drawn to the development of this 3D laser-based data acquisition system and its potential applications. INO/Pavemetrics (Laurent, Lefebvre, and Samson, 2008) developed a commercial system using the line laser imaging technique (fig 2.6). This Laser Crack Measurement System (LCMS) achieves a 0.5 mm resolution on depth, 1 mm resolution on transverse direction, and operates at highway speed (100 km/hr).

The high-cost specialized sensor based methods described above require specialized vehicles and personnel to operate, which increases their cost and makes them difficult


Figure 2.5: Image of pavement at (a) night; (b) daytime with shadows and (c) daytime without shadows (Tsai and Li, 2012)

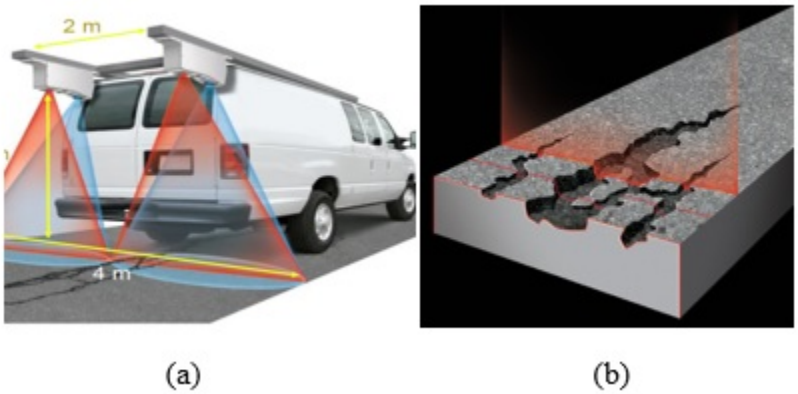



Figure 2.6: Illustration of Pavemetrics LCMS (Laurent, Lefebvre, and Samson, 2008)

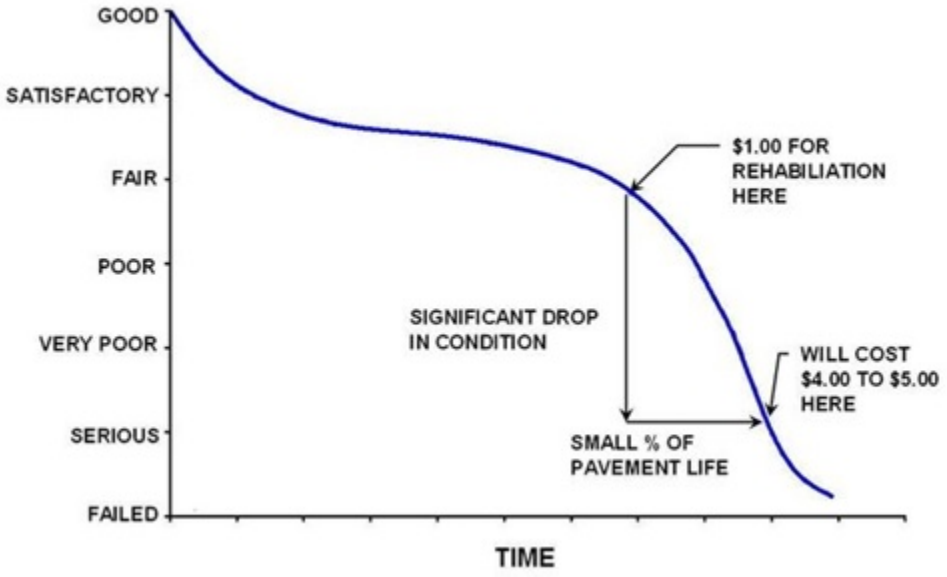



Figure 2.7: Typical deterioration curve of pavements (Federal Aviation Administration, 2017)

to scale. For example, it costs Virginia Department of Transportation (VDOT) USD 1.8 million to survey all interstates, primary roads and 20% of all secondary roads through a contractor using image-based automated surveys (Sauerwein and Smith, 2011). These surveys are carried out only once a year. Haider, Baladi, Chatti, and Dean (2010) show that these image-based pavement condition estimation systems can benefit from more frequent collection of data, and that a higher frequency of data collection will affect the MR&R decision making. Figure 2.7 shows the typical deterioration curve of pavements. As a road segment approaches the precipice in the illustration, more frequent pavement condition assessment is naturally needed to track the drastic change in pavement condition over a short period of time. However, it is not economically feasible to operate laser scanners frequently on all road classifications. Thus, low-cost methods are needed which can be easily scaled for more frequent road condition assessment on all road classifications.

## 2.3 Low-Cost Methods for Functional Performance Measurement

Low-cost vehicle-mounted devices for pavement condition estimation include custom devices which integrate an accelerometer, GPS and data storage and/or transmission capabilities as well as smartphones, which generally have an integrated accelerometer (to measure linear acceleration), gyroscope (to measure angular velocity), magnetometer (to measure orientation/rotation), camera, GPS, data storage and data transmission capabilities. These sensor modules are compact, inexpensive and easy to install. This enables users to crowdsource the pavement condition, rather than have dedicated engineers to collect the pavement condition regularly, leading to a more scalable and frequent system of collecting road condition data. Sauerwein and Smith (2011) compared three methods for crowdsourcing of pavement condition data: using smartphones, using probe vehicles with custom devices, and using connected vehicles with embedded sensors. They concluded that using smartphones was the recommended option among low-cost road condition assessment approaches.

### 2.3.1 Pavement Condition Estimation Using Accelerometer Data

One of the early attempts to crowdsource pavement condition assessment using low-cost vehicle-mounted sensors was by Eriksson, Girod, Hull, Newton, Madden, and Balakrishnan (2008). In their project called *Pothole Patrol*, a combination of accelerometers and GPS sensors on 7 taxis were used to detect over 200 potholes in Boston, MA. The acceleration signal was preprocessed to remove noise and the variance of the acceleration was thresholded to detect potholes. Fixed locations which always caused false positives, such as railroad tracks, were manually blacklisted.

Chen, Lu, Tan, and Wu (2013) detected potholes and road roughness (IRI) using accelerometers and GPS in 100 taxis in Shenzhen, China. The IRI was simply calculated as

a linear function of the variance of the magnitude of acceleration in a road segment.

$$IRI(m/km) = \frac{\sigma - 0.013}{0.5926} \tag{2.1}$$

However, as shown later, the relation between the acceleration captured by a vehicle-mounted accelerometer and the IRI of the road segment depends on a large number of other factors, such as vehicle model, vehicle speed and position of the sensor. Their pothole detection accuracy was 90%, although some fixed location false positives were manually blacklisted. The performance on IRI estimation was evaluated qualitatively. Recently, Aleadelat, Ksaibati, Wright, and Saha (2018) validated the use of acceleration variance for estimating the IRI using a linear transform on tests conducted in 35 road segments in Wyoming. It was found that less than 12% of segments had an error of more than 0.473 m/km (30 in/mile) when estimating the IRI.

Dawkins, Bishop, Powell, and Bevly (2011) conducted a similar study for estimating the IRI of road segments using vehicle-mounted accelerometers. In their study, several approaches for estimating the IRI from the acceleration data were tested. Tests at the National Center for Asphalt Technology (NCAT) test track showed that a pseudo-IRI model gave the lowest RMS error of 0.121 m/km for IRI values smoothed over a 100 m window. However, manual calibration of the accelerometer was needed for every vehicle.

$$IRI = \frac{1}{L} \sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{2.2}$$

Where $x_i$ is the vertical displacement calculated (by double integrating acceleration data) at time step $i$. This pseudo-IRI approach, based on the accumulation of vertical displacements calculated using accelerometer readings, has been widely used as a common approach for IRI estimation using smartphone-collected data (Zang, Shen, Huang, Wan, and Shi, 2018; Islam, Buttlar, Aldunate, and Vavrik, 2014).

Apart from potholes and roughness, low-cost vehicle-mounted devices have also been

used for crowdsourcing other road and traffic parameters. The resonant frequencies of bridges can be used to assess their structural condition. Traditionally this is measured using sensors which have to be embedded directly onto the bridge. Yang and Chang (2009) and Yang, Lin, and Yau (2004) used low-cost vehicle-mounted accelerometers passing over bridges to determine the bridge frequencies. Aleadelat and Ksaibati (2017) explored the prediction of Present Serviceability Index (PSI) using smartphone acceleration. Estimation of raveling has also been explored by measuring the vehicle vibration using smartphones. Using smartphone location in vehicles to assess traffic conditions is an extensively researched topic (Mohan, Padmanabhan, and Ramjee, 2008; Bhoraskar, Vankadhara, Raman, and Kulkarni, 2012; Thiagarajan, Ravindranath, LaCurts, Madden, Balakrishnan, Toledo, and Eriksson, 2009; Vittorio, Rosolino, Teresa, Vittoria, and Vincenzo, 2014; Ghose, Biswas, Bhaumik, Sharma, Pal, and Jha, 2012; Astarita, Caruso, Danieli, Festa, Giofrè, Iuele, and Vaiana, 2012). Vehicle-mounted smartphone cameras have also been recently used to collect pavement images, which can then be used to manually detect distresses such as raveling (Massahi, Ali, Koohifar, Baqersad, and Mohammadafzali, 2018a,b).

Smartphones have also been used on non-motorized vehicles for pavement condition assessment. Zang, Shen, Huang, Wan, and Shi (2018) used bicycle-mounted smartphones to determine the IRI of bike paths by integrating the vertical acceleration captured by the smartphone accelerometers.

$$IRI = \frac{\int \int_{t_{start}}^{t_{stop}} |a_v|(dt)^2}{S} \tag{2.3}$$

Where $a_v$ is the vertical acceleration of the smartphone and $S$ is the length of the road segment. Takahashi, Kobana, Isoyama, Tobe, and Lopez (2017) used smartphones kept with the rider to detects various bumps on bike paths. Complex filters were developed in this study to account for the body movement of the rider themselves on the bicycle.

### 2.3.2 Methods for Crowdsourced Pavement Data Registration and Condition Estimation

Crowdsourcing of pavement condition information is useful only if the collected data can be registered accurately to the road network and the data collected from multiple runs is repeatable. Ndoye, Barker, Krogmeier, and Bullock (2011) show that through the Weak Law of Large Numbers (Papoulis and Pillai, 2002), averaging the sensor responses from multi-run data is an effective method to capture the true response of the pavement only if the data has been properly registered spatially. As shown in figure 2.8b, combining multi-run data without data registration results in loss of information from the runs, instead of removing the noise. Flintsch, Valeri, Katicha, Leon Izeppi, and Medina-Flintsch (2012) tested the feasibility of crowdsourcing pavement condition information using low-cost vehicle-mounted devices by testing the repeatability of these devices. Data registration was done by shifting the series to maximize the cross-correlation between them. They found a high level of repeatability of acceleration readings taken in 3 laps over the same route, with a maximum standard deviation of differences of 0.239 $m/s^2$.

Several methods for registering and combining multiple-run data exist (Kong, Zhao, Wei, and Liu, 2013; Herring, Hofleitner, Abbeel, and Bayen, 2010; Hadachi, Lecomte, Mousset, and Bensrhair, 2011). In practice, the road condition is defined for fixed-length, continuous road segments (GDOT, 2007). The existing methods do not incorporate the fact that the multiple-run data has to be distributed to these discrete road segments during their combination. The existing methods model the road as a linear structure, without consideration that there can be multiple lanes and the vehicles collecting the data can move between lanes, even in the middle of a road segment.

Especially with multiple-run data, the possibility of identifying and removing outliers in the data should be explored to improve the overall performance of the method using the remaining estimates of the pavement condition. Several papers attempt to eliminate outliers by preprocessing the raw data itself (Aihara, Imura, Takasu, Tanaka, and Adachi, 2014; Tang, Yang, Dong, and Li, 2016). Outliers can also be identified and removed by postpro-

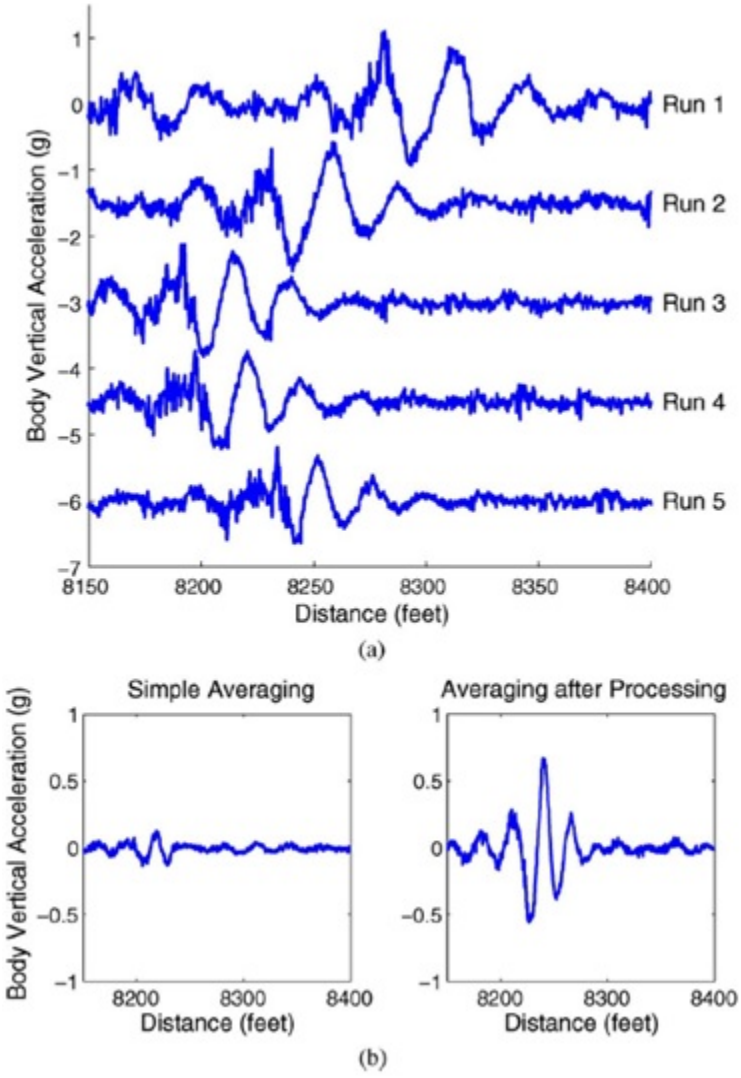



Figure 2.8: Illustration of importance of data registration (a) Example multi-run data before registration; (b) Result of simple averaging before and after data registration (Ndoye, Barker, Krogmeier, and Bullock, 2011)

cessing the obtained pavement condition estimates. Byrne, Isola, and Parry (2013) used the concept of minimum message length (MML) to identify outlier trends from multiple-run data, but requires a measure of the "outlier probability" or confidence level of each estimate as an input. The paper does not specify how to obtain a measure of the confidence level for each estimate.

### 2.3.3 Commercial Products

Researchers from Carnegie Mellon University founded *Roadbotics* (2016) to conduct pavement condition assessment using video data from vehicle-mounted smartphones. Supervised learning models were trained to classify the frames from the collected video data to five levels of pavement condition. Services for data visualization are also provided. *Roadbotics* claims over 100 clients use their product for pavement condition assessment. However, the custom 5-point pavement condition scale used by *Roadbotics* might not be compatible with the requirements of many transportation agencies. For their approach of supervised learning, the collected data will have to be manually relabelled according to the requirements of various transportation agencies to obtain a more useful output. Additionally, some distress measures, such as roughness, is difficult to determine using visual smartphone data. *Roadbotics* employs trained technicians to collect the road condition data. Crowdsourcing of data is currently not available.

Another product for IRI estimation using Android smartphones is *Roadroid*. *Roadroid* simply provides an app for estimating the IRI using the accelerometer sensor data of the smartphone. The parameters of the model used by the app for estimating the IRI have to be manually calibrated for each setup, thus this product is also not ideal for crowdsourcing. *Roadroid* was used by Scholotjes, Visser, and Bennett (2014) for the development of a "smartphone roughness meter" in a World Bank study.

### 2.3.4 Shortcomings and Research Need

One important disadvantage of using mobile devices and crowdsourced vehicles is the sensitivity to user context. The vibration captured by the smartphone for a given road segment can vary by vehicle model, vehicle speed, smartphone position and orientation (Aleadelat and Ksaibati, 2017). Douangphachanh and Oneyama (2013a,b,c) have extensively studied pavement roughness estimation using smartphones in different settings. In their study, smartphones kept in different positions and different vehicles were driven over the same road segment in Vientiane, Laos. It was observed that despite the same road segment being the same, the magnitude of acceleration captured in these road segments differed greatly by smartphone position and vehicle model.

Zeng (2014) required a 6-step calibration process for every new vehicle before it could be used for road condition assessment. Nagayama, Miyajima, Kimura, Shimada, and Fujino (2013) required a similar calibration step for every new vehicle. Zang, Shen, Huang, Wan, and Shi (2018) required riders to maintain a certain speed and bicycle posture. This level of sensitivity on factors which can change with every run (e.g. position and orientation of smartphone inside the vehicle) and even every second (e.g. vehicle speed) makes it difficult to use smartphone data for accurate road condition assessment.

As a result of the user context described above, crowdsourced road condition data leads to a mix of good and bad road condition estimates. However, existing methods do not associate a confidence level with the estimates. This significantly hinders the implementation of smartphone data derived road condition estimation. A confidence level is needed to reduce the effect of outliers when combining estimates for a given road segment. A confidence level associated with the road condition estimates also improves their utility for transportation agencies, which can use the confidence level to prescreen road segments for more testing (Zeng, 2014).

Another shortcoming is that existing literature on low-cost functional performance measurement for road condition assessment focuses only on pavement roughness and potholes.

However, for most transportation agencies, MR&R decision making requires the detection of several other pavement distresses, such as cracking, rutting and raveling (GDOT, 2007; FDOT, 2017), which have not been considered in most of these studies.

Thus, on one hand, mobile devices offer a cost-effective method to frequently gather road condition data on a large scale using crowdsourcing. On the other hand, three major shortcomings have to be addressed: 1) there is no confidence level associated with the condition estimations which diminishes their utility; 2) the existing methods are highly sensitive to the user context, such as vehicle model, vehicle speed, smartphone orientation and position, etc.; and 3) the existing methods focus mainly on just two distresses: pavement roughness and potholes.

## 2.4 Summary

Various methods for evaluating the pavement condition were presented in this literature review. Among the existing approaches, a trade-off between cost and accuracy was found. High-cost performance measurement systems are accurate, but too expensive to scale spatially or temporally. Low-cost approaches are easier to scale, but are limited in their scope and accuracy. The following research needs have been identified:

1. The low cost and ubiquity of smartphones offer an effective solution for cost effective and frequent pavement condition assessment. However, the following shortcomings have made it difficult for transportation agencies to apply smartphone based pavement condition assessment solutions:

    (a) Existing research on pavement condition assessment using smartphone sensor data has focused on pavement roughness (IRI) and the presence of potholes. The application of smartphone based pavement condition assessment in the context of important distresses such as cracking, rutting and raveling has not been adequately explored. Thus, the possibility of detecting and classifying

pavement distresses other than roughness and potholes needs to be explored.

(b) The response from smartphone sensors for a given road segment can vary greatly due to factors other than the pavement condition, such as vehicle model, vehicle speed, smartphone position and smartphone orientation. This sensitivity to user context has caused existing solutions to be extremely limited in scope (specific vehicle needed, specific smartphone position needed) or require extensive calibration before usage for pavement condition assessment. Thus, an accurate method for pavement condition assessment using smartphone data without the need for extensive calibration or additional input is needed.

(c) Existing methods for pavement estimation using smartphone data do not associate a level of confidence with the estimations. This presents two issues:

   i. Crowdsourced pavement condition estimations for a given road segment have to be processed to provide a combined pavement condition estimation. Without a confidence level for each estimation, a mix of good and poor pavement condition estimations are combined together with equal weightage, lowering the accuracy of the combined pavement condition estimations due to outliers.

   ii. A confidence level for an estimate increases its utility. For example, there is a 30% chance of rain tomorrow conveys more useful information than it may rain tomorrow. In the context of pavement condition assessment, engineers can use the confidence level of the pavement condition estimations to shortlist only low confidence road segments for further investigation, reducing effort (Zeng, 2014). However, in existing research, without such a confidence level, there is no method to identify the road segments which require further investigation.

Thus, a method for smartphone based pavement condition estimation with an associated confidence level is strongly needed.

2. 3D laser technology has become a mainstream technology for pavement condition assessment because of its high accuracy and robustness against environmental conditions. However, because of its high cost, 3D laser based pavement condition estimation is used for only a subset of roads and at low frequency (annually or bi-annually). This limitation leads to the following problems:

   (a) 3D laser based pavement condition assessment is generally limited to interstates and highways only. It may not be feasible to survey all types of roads due to economic constraints. Crowdsourcing of pavement condition data is a scalable solution which can be used to augment pavement condition assessment using 3D laser technology. However, a method to register multi-run data from different types of sensors for the purposes of pavement condition estimation is needed.

   (b) For road segments with severe deterioration of the pavement condition, the pavement condition worsens rapidly. Thus, frequent pavement condition assessment (monthly, quarterly) is needed for timely treatment of the affected road segments. Unfortunately, frequent pavement condition estimation using 3D laser technology is not economically feasible, especially for smaller transportation agencies (city-level or county-level). Crowdsourced road condition data using low-cost sensors such as smartphones can potentially help to improve pavement condition forecasting between pavement condition assessments using 3D pavement data. Thus, the possibility of using low-cost pavement condition assessment methods, including smartphone based methods, to enhance the pavement condition forecasting of 3D pavement data needs to be explored.

This thesis aims to combine smartphone data and 3D pavement data to create a hybrid approach to pavement condition estimation and forecasting that overcomes the disadvantages of the individual approaches. Chapter 3 addresses research need 2a. In chapter 4, a

methodology for single-run pavement condition estimation using smartphone data is presented which attempts to address research needs 1a and 1b. In chapter 5, methodologies for multiple-run pavement condition estimation and forecasting are presented to address the research needs in point 1c and 2b.

# CHAPTER 3
# METHODOLOGY FOR SMARTPHONE AND 3D PAVEMENT DATA COLLECTION AND REGISTRATION

This chapter explains the methodology to collect and register smartphone and 3D pavement data from multiple runs of data collection. First, the procedure for data collection using smartphones and 3D pavement data is explained. Second, a Geographic Information System (GIS) model of the road network to which the collected data has to be registered is developed. Third, a method for registering smartphone and 3D pavement data onto this GIS model is proposed. Finally, the proposed methodology is validated through case studies.

## 3.1 Data Collection Procedure

Smartphone and 3D pavement data were collected using on-vehicle surveys. In this thesis, a "run" refers to a continuous recording of data during an on-vehicle survey using the sensing van, smartphones or both. A "scenario" represents a certain set of conditions during a run. One run can collect data for multiple scenarios, for example different phones collecting data simultaneously. The following subsections explain the 3D pavement and smartphone data collection procedures.

### 3.1.1 3D Pavement Data Collection Procedure

3D pavement data was collected using the Georgia Tech Sensing Vehicle (GTSV). The GTSV is equipped with two laser scanners (Pavemetrics Laser Crack Measurement System) (fig. 2.6), two LIDAR sensors, four cameras, GPS, and an inertial measurement unit (IMU). The laser scanner collects both the intensity of the signal reflected from the pavement surface and the range (3D pavement data), as shown in figure 2.4. The range image is particularly robust against changing lighting conditions and shadows, which adversely

impact camera images (fig 2.5). The range data consists of points collected 1 mm apart transversely and 5 mm apart longitudinally on the pavement surface. 4,160 points are collected in each transverse profile, thus covering slightly more than one lane of roadway. Typically 1,000 transverse profiles are stored together as an individual image, covering a 4,160 mm $\times$ 5,000 mm pavement surface area.

### 3.1.2 Smartphone Data Collection Procedure

I had developed an Android app called AllGather to record sensor data from smartphones. AllGather simultaneously collects the signals from the smartphone camera, accelerometer, magnetometer, gyroscope and GPS. Although this study primarily focused on the pavement surface which can be assessed from the vibration response, the image data can be used to assess the condition of additional road infrastructure components, such as guardrails and traffic signs. The Xiaomi 8 and LG G5 smartphones were used to collect data, though any Android smartphone with the required sensors and OS Android 5.0 or above is supported. The app is also available on iOS.

The phone was mounted on a car holder on the windshield as shown in figure 3.1. The app collected video at $1920 \times 1080$ resolution at 30 frames per second of the view from the windshield. GPS points were collected at a frequency of 1 Hz. Sensor data (accelerometer, magnetometer and gyroscope) were collected at a rate of 50 Hz.

With these settings, AllGather collected data at 10 Mbps (approximately 1 GB of data collected every 13 minutes). The bulk of the collected was the video data (99.8%). The data rate can be reduced by lowering the image resolution or not collecting video data altogether. One of the smartphone models (LG G5) faced temperature issues during the Summer. Power consumption and temperature can also be controlled by turning off the video data collection, making the smartphone data collection procedure suitable for crowdsourcing.

Figure 3.1: Smartphone data collection app

### 3.2 Proposed Methodology for Data Registration

In practice, the road condition is defined for fixed-length, continuous road segments (GDOT, 2007). A single data collection run (using either laser scanners or smartphones) can go through multiple road segments with different road conditions. For 3D pavement data, a data point refers to one image of the pavement surface and for smartphone data, a data point refers to one record of accelerometer, magnetometer and gyroscope data. Collected 3D pavement and smartphone data have location information attached to each data point collected. Therefore, after a run is collected, the collected data points must be split to correspond to predefined discrete road segments, for which these data points provide the pavement condition.

The road conditions from smaller road segments can be aggregated to estimate the condition of a larger segment. However, the road segment length should not be so small that

minor anomalies on the pavement surface can significantly affect the overall road segment condition. In this thesis, 16.4 feet (5 meters) was used as the standard length for the road segments to correspond to the length of the intervals for which the labelled ground truth was provided from the laser scanner. The ARCMap software was used to create a GIS model of the data collection route (figure 6.6) and to break the route into 5-meter road segments. Each road segment is defined by the chainage of its starting point along that route. This concept can be easily extended to multiple routes as well. Thus, given the chainage of the starting point of each segment and the length of each segment (5 meters), the collected data can be registered to the segments if the chainage of each data point along the route is known. The following subsections explain the proposed methodology for determining the chainage of each data point of 3D pavement and smartphone data.

### 3.2.1 Data Registration for 3D Pavement Data

Data registration for 3D pavement surface data is straightforward. The GTSV uses a distance measurement unit to trigger data collection, which ensures that the laser scanner collects a continuous scan of the pavement surface without overlap at a fixed distance interval. The collected images are split into images covering 16.4 feet (5 meters) of distance travelled, and each image has an associated GPS location. The location of the starting point of the first road segment is known from the GIS model. The closest 3D pavement data point $i$ to this starting position was first determined, which is assigned the initial chainage (in meters) $c_i$ (typically 0). The chainage (in meters) of any subsequent data point $j$ can be determined simply as

$$c_j = 5(j - i) + c_i \tag{3.1}$$

Thus, using linear referencing, each 3D pavement data point is registered onto the road network.

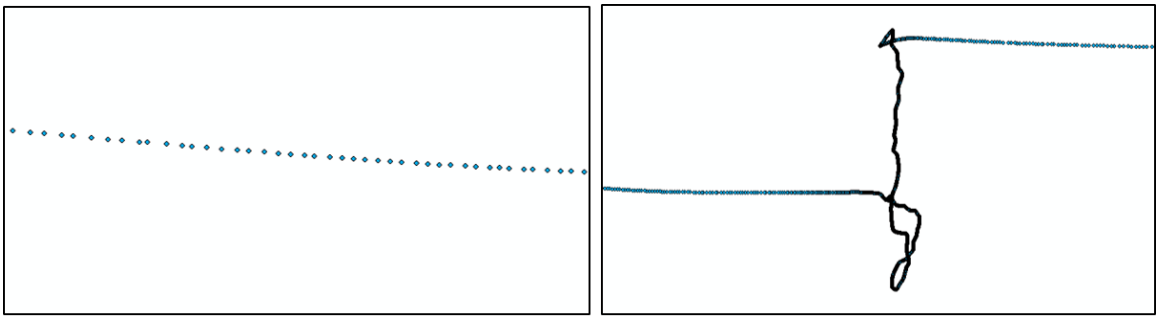
### 3.2.2 Data Registration for Smartphone Data

For smartphones, data registration is more complex. The sensor data is collected at fixed frequencies, not distances. Moreover, because of power consumption constraints, GPS data is generally collected at a lower frequency than the data points from accelerometers, magnetometers and gyroscopes. As described in section 3.1.2, GPS data was collected at a frequency of 1 Hz, while sensor data was collected at a higher frequency of 50 Hz, which are typical frequencies for modern smartphones. To register the smartphone data, the following method is proposed:

1. Location data is typically collected in decimal degrees latitude and longitude. For the mathematical operations required for registering the smartphone data points, the location data has to be first projected onto a projected coordinate system where the location coordinates are part of a Cartesian plane. The Universal Transverse Mercator (UTM) Zone 16N projected coordinate system is a suitable projected coordinate system for coordinates near Atlanta, GA, USA, where the data for this thesis was collected. Thus, the smartphone location data was first converted to the UTM Zone 16N projected coordinate system coordinates.

2. After converting the smartphone location data to the projected coordinate system, the location of each data point can be interpolated. The piecewise cubic hermite interpolant polynomial (PCHIP) method is used to interpolate the projected coordinates at the timestamp of each sensor data point (Ndoye, Barker, Krogmeier, and Bullock, 2011; Fritsch and Carlson, 1980). After interpolation, each data point has associated projected coordinates. However, because the smartphone data points are collected at a fixed time frequency, the spatial distance between these data points is not constant (figure 3.2a) and can erroneously show movement even when the vehicle has stopped (figure 3.2b).

3. Finally, the chainage of each data point is determined using the following steps. The



(a) (b)

Figure 3.2: Interpolated location of smartphone data points showing (a) uneven spacing and (b) erroneous movements during stops

algorithm for chainage assignment is formally shown in algorithm 1. Figure 3.3 illustrates each step of the chainage assignment.

(a) The predefined road segments are first defined as a series of points $s_j \forall j \in \{1, 2, ...n\}$ where segment $j$ connects $s_j$ and $s_{j+1}$. The location and chainage for each point $s_j$ is known from the GIS model. Thus, point $s_j$ has projected coordinates $\mathbf{x}_s^{(j)}$ and chainage $c_s^{(j)}$.

(b) A 2D-tree of the points $s_j$ is constructed with their projected coordinates $\mathbf{x}_s^{(j)}$ as the key and the index of the point $j$ as the value.

(c) For each data point $p_i$ with interpolated projected coordinates $\mathbf{x}_p^{(i)}$, the closest point on the road network $b$ is found efficiently using the 2D-tree (figure 3.3a). Thus, the point $i$ will have a chainage $c_p^{(i)}$ between $c_s^{(b-1)}$ and $c_s^{(b+1)}$.

(d) To determine whether the point $\mathbf{x}_p^{(i)}$ should be projected onto segment $b-1$ or $b$, the angle between the vector joining $\mathbf{x}_s^{(b)}$ and $\mathbf{x}_p^{(i)}$ and the two segments $b-1$ and $b$ are determined (figure 3.3b). $\mathbf{x}_p^{(i)}$ is projected onto the segment with which it makes the smaller angle (figure 3.3c).

(e) The chainage of data point $p_i$ is then interpolated using the projected vector. If the projection is negative, which occurs if both $b-1$ and $b$ form obtuse angles, then $c_s^{(b)}$ is taken as the chainage of data point $p_i$.

(f) As shown in figure 3.2b, the points can erroneously wander even when the vehicle is stopped. It is assumed that the vehicle does not reverse on the route. Hence, it is ensured that the chainage of each data point $c_p^{(i)}$ is greater or equal to $c_p^{(i-1)}$. If not, then $c_p^{(i)}$ is set to $c_p^{(i-1)}$.

**Algorithm 1** Smartphone Sensor Data Point Chainage Calculation

Let $s^{(i)} \forall i \in \{1, ..., n_s + 1\}$ be the series of points defining the $n_s$ road segments such that segment $i$ spans from $s^{(i)}$ to $s^{(i+1)}$.
Let $\mathbf{x}_s^{(i)}$ be the coordinates of $s_i$ in UTM 16N projected coordinate system.
Let $c_s^{(i)}$ be the chainage of $s_i$.
Let $\mathbf{x}_p^{(i)} \forall i \in \{1, ..., n_p\}$ be the coordinates of the $i$th sensor data point in UTM 16N projected coordinate system.
**for** $i \in \{1, ..., n_p\}$ **do**
    $b \leftarrow argmin_j |\mathbf{x}_p^{(i)} - \mathbf{x}_s^{(j)}|$ ▷ Find the nearest road segment point to the current sensor data point
    $\mathbf{v}_{pb} \leftarrow \mathbf{x}_p^{(i)} - \mathbf{x}_s^{(b)}$
    **if** $b = 1$ **then**
        $d \leftarrow 2$
    **else if** $b = n_s + 1$ **then**
        $d \leftarrow n_s$
    **else**
        $\mathbf{v}_{ab} \leftarrow \mathbf{x}_s^{(i-1)} - \mathbf{x}_s^{(b)}$
        $\mathbf{v}_{cb} \leftarrow \mathbf{x}_s^{(i+1)} - \mathbf{x}_s^{(b)}$
        **if** $(\mathbf{v}_{pb} \cdot \mathbf{v}_{ab})/(|\mathbf{v}_{pb}||\mathbf{v}_{ab}|) > (\mathbf{v}_{pb} \cdot \mathbf{v}_{cb})/(|\mathbf{v}_{pb}||\mathbf{v}_{cb}|)$ **then** ▷ Take the side with which the point makes a smaller angle
            $d \leftarrow a$
        **else**
            $d \leftarrow c$
        $\mathbf{v}_{db} \leftarrow \mathbf{x}_s^d - \mathbf{x}_s^b$
        $t \leftarrow (\mathbf{v}_{pb} \cdot \mathbf{v}_{db})/(|\mathbf{v}_{db}|^2)$
        **if** $t < 0$ **then**
            $c_p^{(i)} \leftarrow c_s^{(b)}$
        **else**
            $c_p^{(i)} \leftarrow c_s^{(b)} + t(c_s^{(d)} - c_s^{(b)})$
**for** $i \in \{2, ..., n_p\}$ **do** ▷ Ensure chainage is monotonically increasing
    $c_p^{(i)} \leftarrow max(c_p^{(i)}, c_p^{(i-1)})$

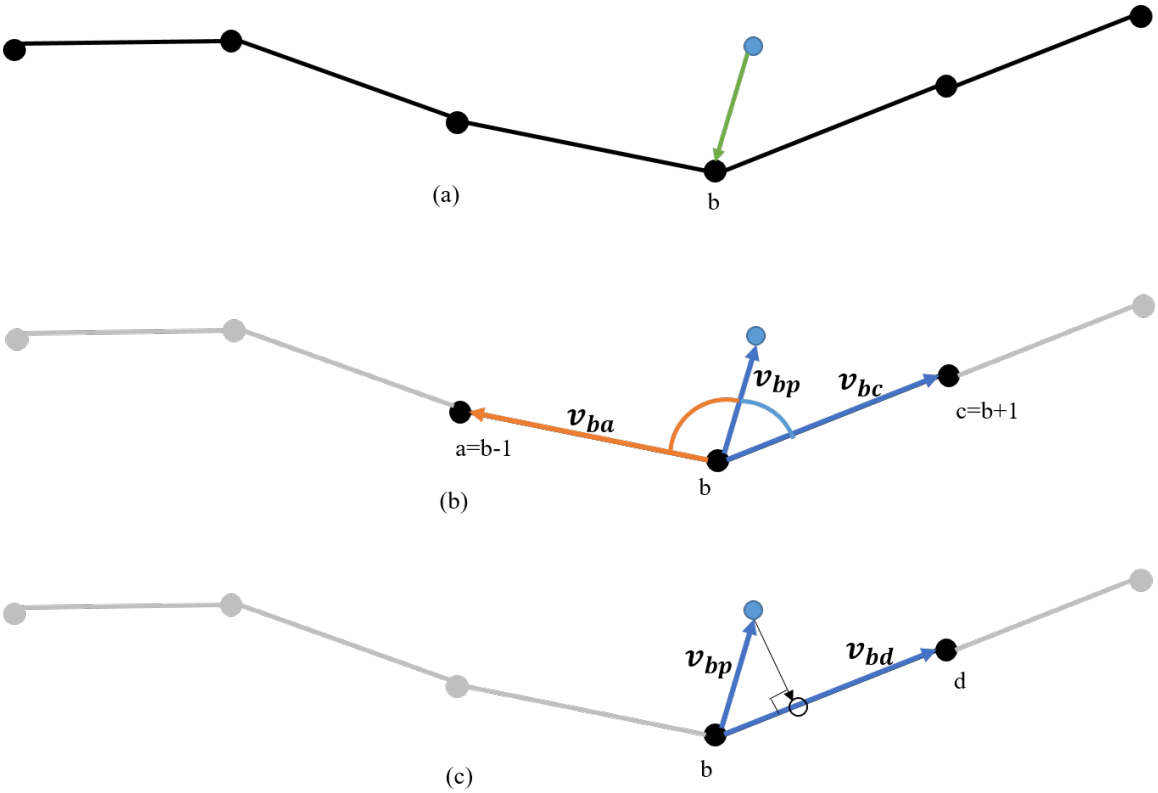



Figure 3.3: Illustration of smartphone data point chainage calculation (a) search for nearest road segment point $b$, (b) choose adjacent segment with lower angle difference, (c) project data point onto segment and interpolate chainage

### 3.3 Validation of Proposed Methodology for Data Registration

Pearson cross correlation between two series of data is a measure of the similarity between the two series. Higher Pearson correlation indicates higher similarity between the two series. The Pearson correlation $(x * y)[k]$ between two series $\{x_1, x_2, ....x_n\}$ and $\{y_1, y_2, ..., y_n\}$ for a given lag between the two series $k$ is defined as

$$(x * y)[k] = \frac{\sigma_{xy}(k)}{\sqrt{\sigma_{xx}(0)\sigma_{yy}(0)}} \tag{3.2}$$

Where

$$\sigma_{a,b}(k) = \frac{1}{n}\sum_{t=1}^{n}(a_{t-k} - \bar{a})(b_t - \bar{b}) \tag{3.3}$$

Where

$$\bar{a} = \frac{1}{n}\sum_{t=1}^{n} a_t \tag{3.4}$$

And

$$\bar{b} = \frac{1}{n}\sum_{t=1}^{n} b_t \tag{3.5}$$

If the 3D pavement and smartphone data is successfully registered, the series of distress values calculated by each type of data from different runs on the same route will be highly cross-correlated at zero lag ($k = 0$), with the cross-correlation dropping on either side if a lag between the two series is introduced.

To validate the proposed methodology for data registration, 3D pavement and smartphone data were repeatedly collected for the data collection route, as described in section 6.2. The distress values for IRI were calculated for both the 3D pavement and smartphone data as explained in chapter 4. The following subsections describe the validation of 3D pavement smartphone data registration using the principle explained in this section.

### 3.3.1 Validation of Proposed Data Registration Methodology for 3D Pavement Data

As mentioned in section 3.2.1, the 3D pavement data is collected at a constant distance interval of 5 meters. However, there can be an offset between each run, which must be fixed by data registration to correspond each run to the predefined road segments. The IRI values calculated without data registration are plotted from each individual run in figure 3.4. A shift can be observed between the IRI estimated from each run because of the lack of registration. After following the data registration method proposed in section 3.2.1, the corresponding plot of distress values from each individual run are given in figure 3.5. It can be observed that the proposed data registration method successfully minimizes the shift between the individual runs.

To quantitatively validate the proposed 3D pavement data registration method, the cross-correlation between the distress values from two different runs can be plotted for

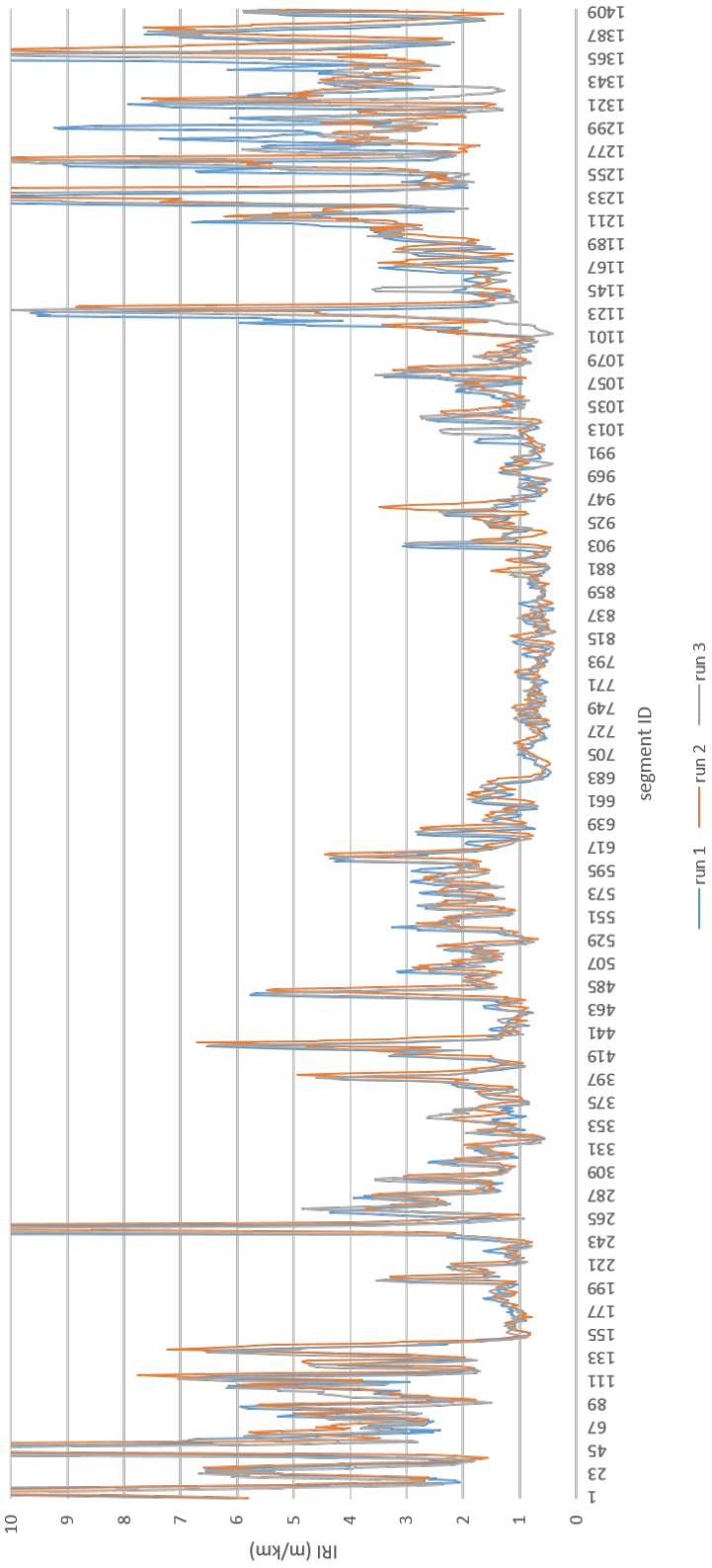



Figure 3.4: IRI from multiple run 3D pavement data before data registration

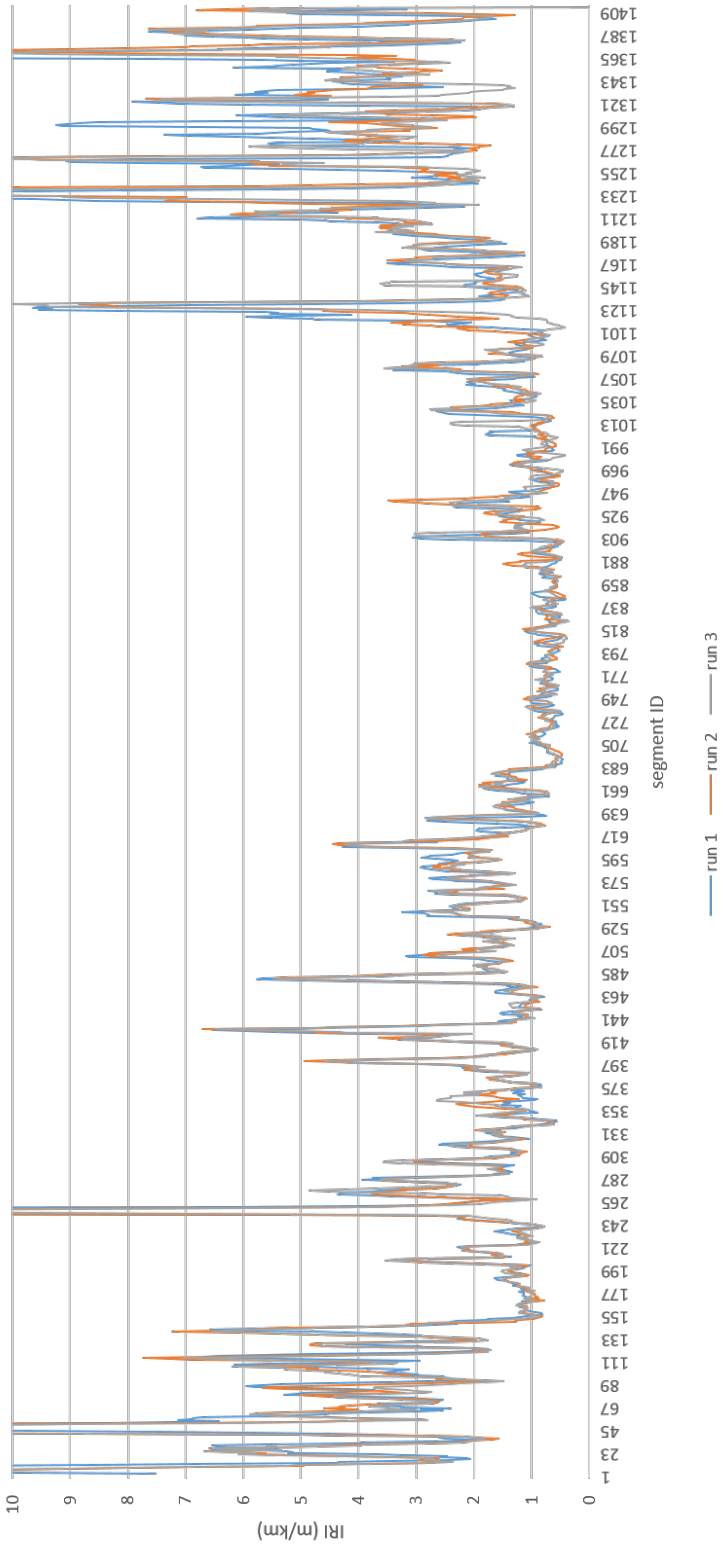



Figure 3.5: IRI from multiple run 3D pavement data after data registration

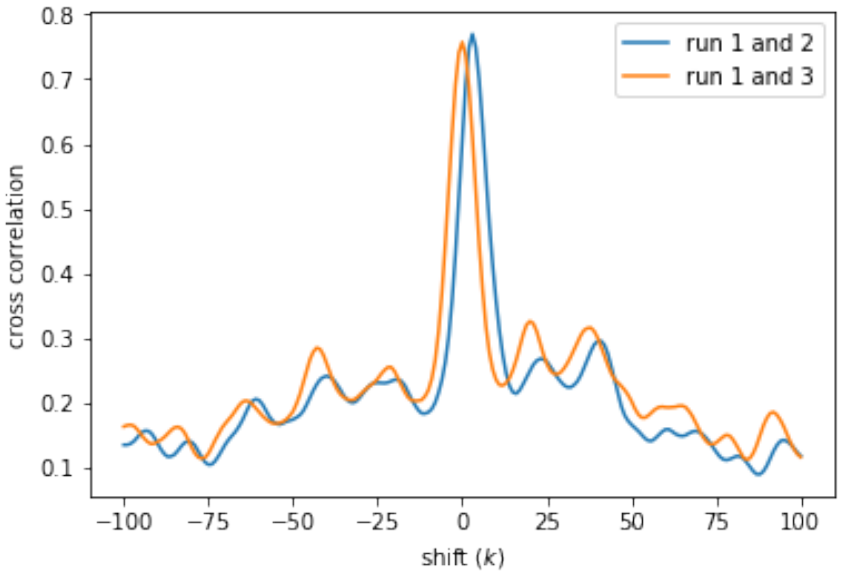



Figure 3.6: Cross-correlation between IRI values before data registration

various shifts ($k$) between the two series, as shown in figure 3.6. It can be observed that run B was started 75 m before run A and run C was started 5 m before run A.

After data registration using the method presented in section 3.2.1, the cross-correlation plots are shown in figure 3.7. A high cross-correlation is observed when there is no shift ($k = 0$) between any of the runs. Any further shift would introduce a significant drop in the cross-correlation between the runs, indicating that the proposed method successfully registered the 3D pavement data.

### 3.3.2 Validation of Proposed Data Registration Methodology for Smartphone Data

For smartphone data, it is not possible to calculate and visualize the distress values before data registration since the smartphone data points are not collected at a fixed distance interval. Chainage for the smartphone data points are only assigned after the methodology presented in section 3.2.2 is completed.

After following the data registration method proposed in section 3.2.2, the corresponding plot of IRI distress values from each individual run are given in figure 3.8. The dashed vertical lines are drawn to indicate the correspondence between peaks in the curves from

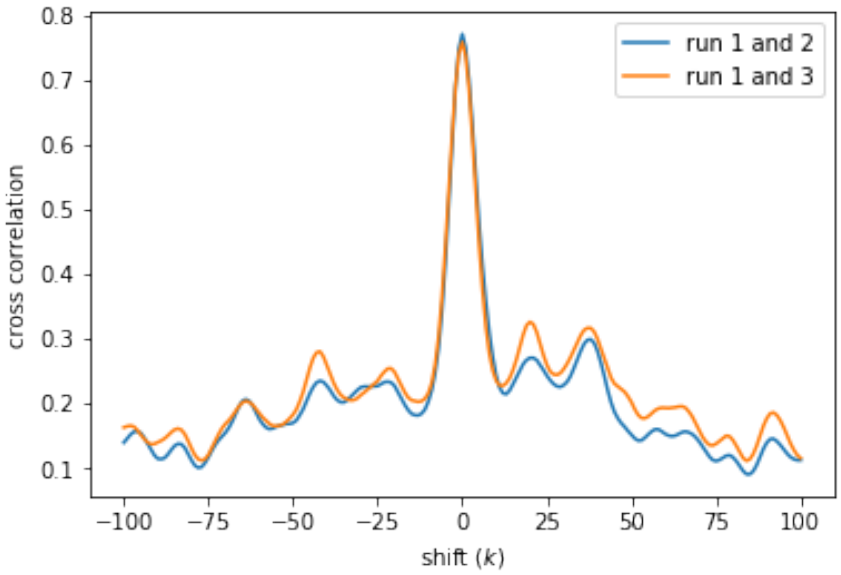



Figure 3.7: Cross-correlation between IRI values after data registration

each individual run, indicating a successful registration. The peaks in IRI values generally occur at intersections, where pavement roughness is particularly high.

To quantitatively validate the proposed smartphone data registration method, the cross-correlation between the distress values from two different runs after data registration were plotted for various shifts ($k$) between the two series (figure 3.9). Again, a high cross-correlation is observed when there is no shift ($k = 0$) between any of the runs. Any further shift would introduce a significant drop in the cross-correlation between the runs, indicating that the proposed method successfully registered the smartphone data.

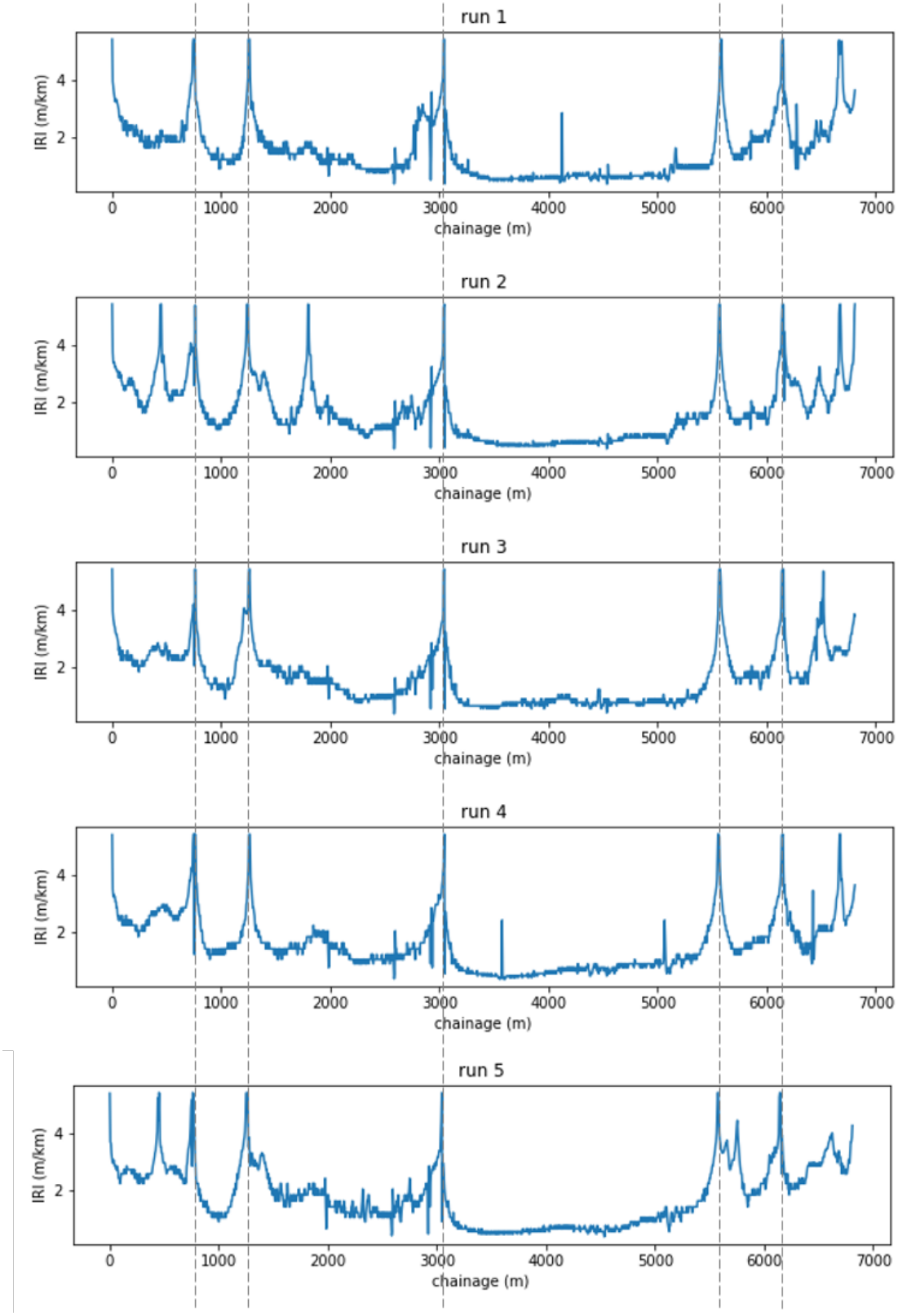



Figure 3.8: IRI from multiple run smartphone data after data registration

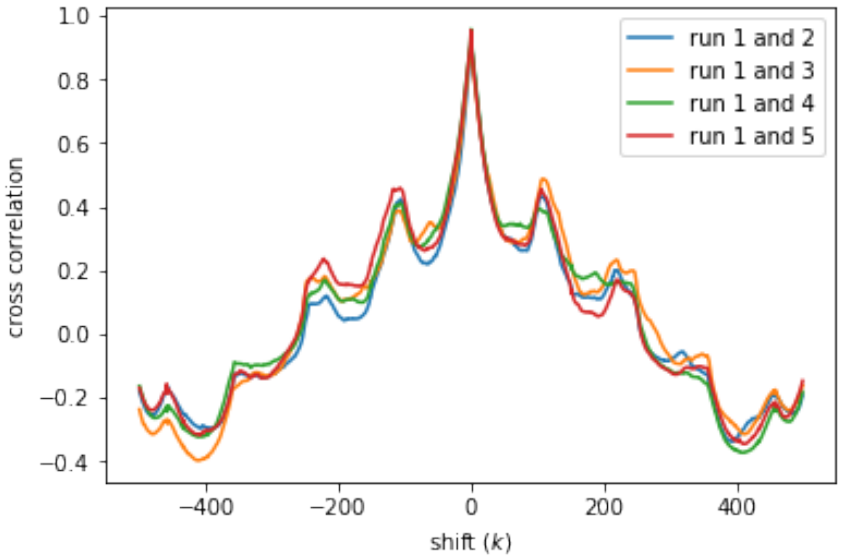



Figure 3.9: Cross-correlation between IRI values from smartphone after data registration

# CHAPTER 4
# METHODOLOGY FOR SINGLE RUN PAVEMENT CONDITION ESTIMATION USING SMARTPHONE AND 3D PAVEMENT DATA

This chapter explains the methodology to estimate the pavement condition by integrating high-frequency smartphone sensor data and low-frequency high-accuracy 3D pavement data. 3D pavement data is used to provide labelled pavement condition information for training a model to estimate the pavement condition using a single run of smartphone data over a road segment.

First, the proposed methodology for training a model for single-run pavement condition estimation is explained. This is followed by the training results and validation of the proposed methodology.

## 4.1 Proposed Methodology for Single-Run Pavement Condition Estimation

### 4.1.1 Background

In supervised learning, a parameterized model $f_\theta$ which can map the values from an input domain $\mathbb{X}$ to an output range $\mathbb{Y}$ ($f_\theta : \mathbb{X} \rightarrow \mathbb{Y}$) is trained to map a provided input $x \in \mathbb{X}$ to the desired output (label) $y \in \mathbb{Y}$ using provided input-label pairs. $f_\theta$ is initialized with random values for the parameters $\theta$. Given training input-label examples $(x^i, y^i) \forall i \in 1, ..., n$, the parameters are adjusted to make the model more likely to output the given desired label on receiving the corresponding input. Regularization of the model constrains it to help it generalize better for input values which were not included in the training set. In stochastic models, the output is modeled as a random variable, and the model estimates the probability distribution of the output variable. The performance of the trained model is then evaluated on a set of test input-label examples separate from the set of training examples.

Thus, for model training, the input features, corresponding labels, parameterized model and a method for optimizing the parameters of the model (optimizer) is needed. In our specific problem of estimating the pavement condition using smartphone data, the input features are provided by the smartphone sensors and the labels represent the ground truth pavement condition. Thus, the successfully trained model should take as input the smartphone sensor data from a single run over a road segment and output an estimate of the pavement condition. The following subsections explain each component in detail.

### 4.1.2 Input Features

Following the methodology presented in chapter 3, registered smartphone acceleration, magnetometer and gyroscope data is available as input for estimating the pavement condition. For a single run, a given road segment $j$ which spans from chainage $c_s^j$ to $c_s^{(j+1)}$ contains a series of data points $\{p_i\} \forall c_p^{(i)} \in \left[c_s^{(j)}, c_s^{(j+1)}\right)$ collected using a vehicle-mounted smartphone.

The smartphone data points are collected at a fixed time frequency. Thus, depending on the speed of the vehicle, the length of the series for a road segment can vary. Figure 4.1 shows the distribution of the lengths of these series from three different runs of the data collection route. Clearly it is a highly skewed distribution with a small number of segments having series of lengths of up to 4000 data points. Such a high number of data points can be collected on a single road segment when the vehicle comes to a stop (for example at an intersection). Most of the data points in these long series are not useful for pavement condition estimation because the data being collected does not describe the response of the vehicle driving over the road. Thus, the series length can be clamped to a maximum value without a significant loss of useful data points. Figure 4.2 shows the distribution of series lengths for lengths smaller than 100 data points only. To reduce the computation effort, the series lengths for each road segment was capped at 100 data points. This only affected 0.7% of the road segments.

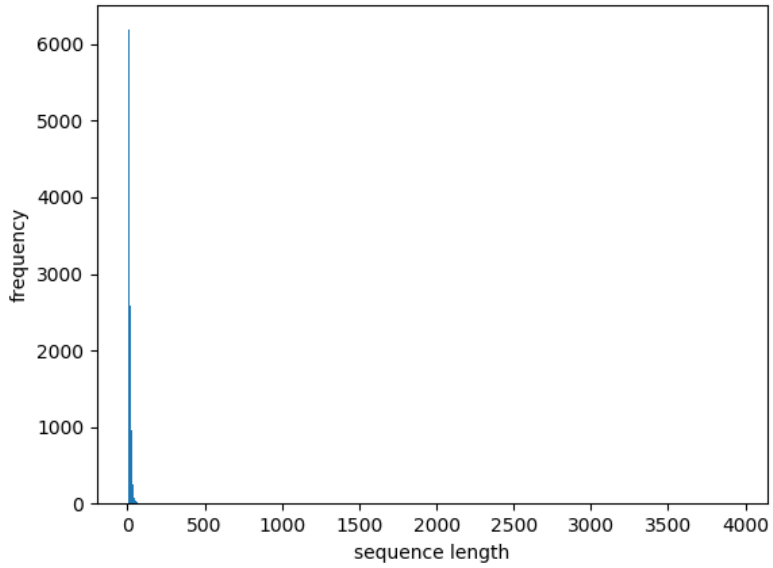



Figure 4.1: Distribution of length of smartphone data point series in road segments

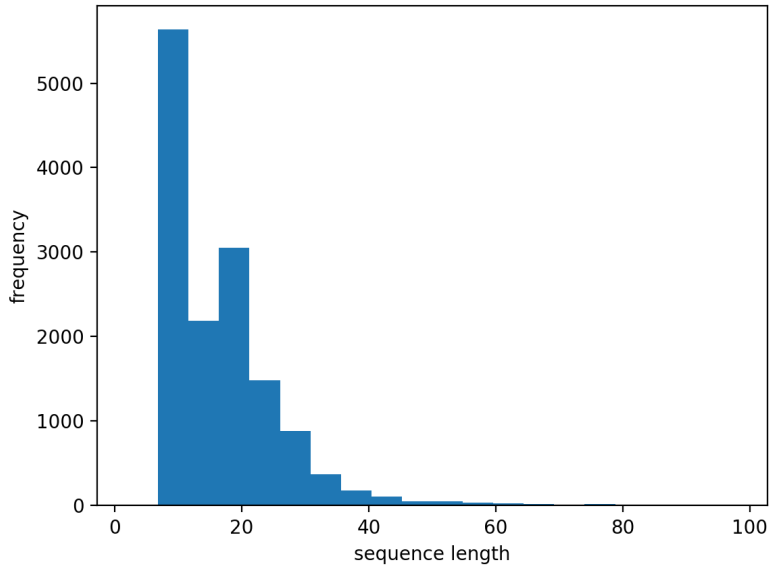



Figure 4.2: Distribution of length of smartphone data point series in road segments for lengths $< 100$ data points

The specific features in each data point is listed in table 4.1. Note that the location information (projected coordinate location, chainage) for the data point was deliberately omitted to prevent the model from learning the pavement condition based on location. However, the speed of the vehicle is a useful feature which cannot be calculated without the chainage. Thus, the average speed is the only handcrafted feature which was added to the raw smartphone sensor data. The average speed of the vehicle $v_j$ in road segment $s_j$ was calculated as

$$v_j = \frac{c_p^{(b)} - c_p^{(a)}}{t_b - t_a} \tag{4.1}$$

Where $a$ and $b$ are the indices of the first and last data points in road segment $j$ respectively and $t_a$ and $t_b$ are the timestamps when those data points were collected. It is expected that the model will learn how to use the provided speed information to minimize the distress value estimation error.

The frame of reference for linear acceleration and angular velocity is aligned with the smartphone. The orientation of the axes for Android devices is given in figure 4.3. The frame of reference for orientation is based on the location of the smartphone in the world, as defined by the Android operating system: the $+z$ axis points upwards away from the ground, the $+y$ axis is parallel to the ground pointing towards the magnetic north pole and the $+x$ axis is given by the outer product of the $+y$ and $+z$ axes, thus parallel to the ground pointing eastward.

### 4.1.3 Ground Truth Labels

As explained in chapter 3, the 3D pavement data is collected at a fixed distance interval of 5 meters. Thus, each road segment contains exactly one 3D pavement data point. The structure of the raw 3D pavement data has been described in section 3.1.1. The methodology presented in section 3.2.1 was followed to register the 3D pavement data to each road segment.

Table 4.1: Features of a smartphone data point

| Feature Name | Source | Description |
|---|---|---|
| `timestamp` | system clock | timestamp of data point |
| `lin_accel_x` | accelerometer | Component of linear acceleration of the phone along $+x$ axis in $m/s^2$ |
| `lin_accel_y` | accelerometer | Component of linear acceleration of the phone along $+y$ axis in $m/s^2$ |
| `lin_accel_z` | accelerometer | Component of linear acceleration of the phone along $+z$ axis in $m/s^2$ |
| `rot_x` | magnetometer | $u_x \sin(\theta/2)$ component of unit quaternion defining orientation of the phone |
| `rot_y` | magnetometer | $u_y \sin(\theta/2)$ component of unit quaternion defining orientation of the phone |
| `rot_z` | magnetometer | $u_z \sin(\theta/2)$ component of unit quaternion defining orientation of the phone |
| `rot_u` | magnetometer | $\cos(\theta/2)$ component of unit quaternion defining orientation of the phone |
| `angvel_x` | gyroscope | Component of angular velocity of the phone along $+x$ axis in $radians/s$ |
| `angvel_y` | gyroscope | Component of angular velocity of the phone along $+y$ axis in $radians/s$ |
| `angvel_z` | gyroscope | Component of angular velocity of the phone along $+z$ axis in $radians/s$ |
| `avg_speed` | calculated | Average speed of vehicle in the current road segment in mph |

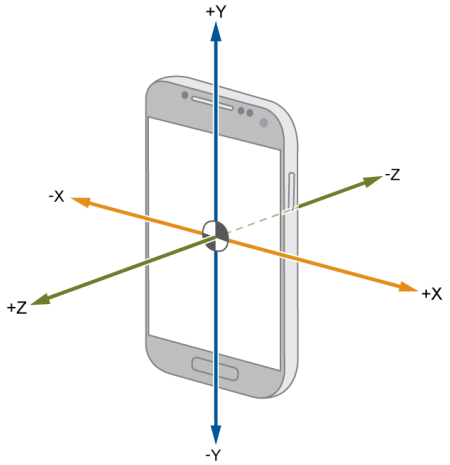



Figure 4.3: Orientation of axes for smartphone accelerometer and gyroscope

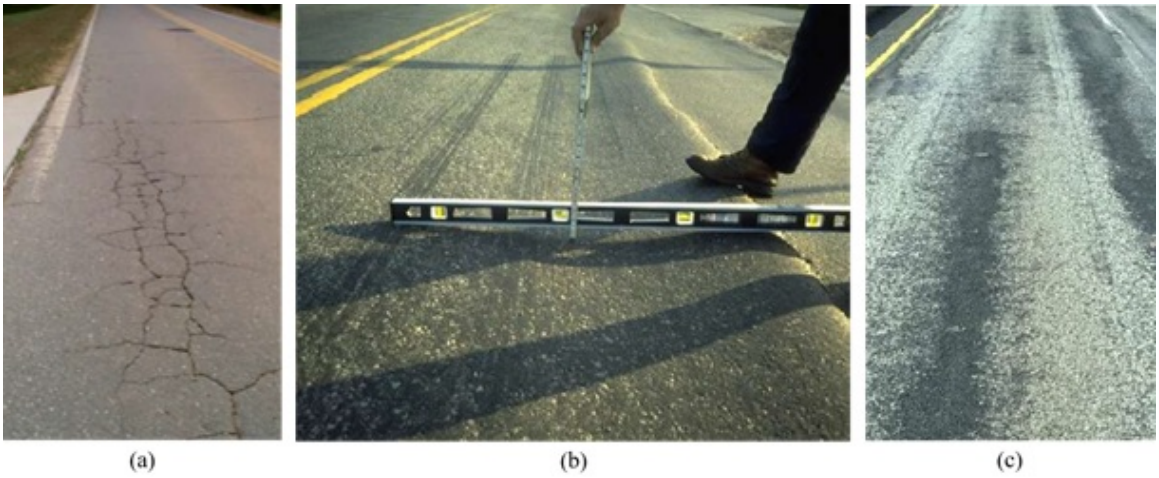



Figure 4.4: (a) cracking, (b) rutting and (c) raveling on asphalt pavements (GDOT, 2007)

Existing literature on pavement condition estimation using low-cost vehicle mounted sensors (including smartphones) focus mainly on measuring roughness using IRI or PSD (Dawkins, Bishop, Powell, and Bevly, 2011; Sauerwein and Smith, 2011; Flintsch, Valeri, Katicha, Leon Izeppi, and Medina-Flintsch, 2012; Zang, Shen, Huang, Wan, and Shi, 2018; Zeng, Park, Smith, and Parkany, 2018; Robinson, 2012). Detection of potholes using smartphones is also commonly explored (Eriksson, Girod, Hull, Newton, Madden, and Balakrishnan, 2008; Chen, Lu, Tan, and Wu, 2013). IRI is a popular method for quantifying the functional pavement condition (FDOT, 2017; Sun, 2003). In this thesis, IRI is used as the primary measure for quantifying the pavement condition.

However, the laser scanner also provides detailed information on various surface distresses such as cracking, rutting and raveling (figure 4.4). The presence and severity of these distresses significantly affect the road condition (GDOT, 2007; FDOT, 2017) and MR&R decision making. Thus, the 3D pavement surface data can help to explore the possibility of detecting these distresses in addition to roughness and potholes using smartphone data. Rutting, raveling, cracking and potholes will also be quantified using the 3D pavement data and the estimation of these distresses using smartphone data will also be explored.

For uniformity, each distress was represented by a real value representing its severity, referred to as a distress value. For roughness, the IRI (in m/km) was used. For rutting, the average rut depth (in mm) was used, which is used for quantifying the pavement condition by GDOT (GDOT, 2007). Many state DOTs use a subjective categorical scale (low, medium, high) to measure raveling, where the definition varies by state to state. Thus, the raveling index (Laurent, Hébert, Lefebvre, and Savard, 2012) was used to represent the raveling severity, which can then be used to determine the raveling category according to different states' needs. Only the number of potholes in a road segment is generally needed for maintenance purposes (GDOT, 2007; Miller and Bellinger, 2014). For potholes, the softmax probability of pothole presence in a road segment is used. Multiple potholes in the same 5 meter road segment are rare, and when the do occur, create a vibration response where the number of potholes is very difficult to distinguish, even manually. Thus, the number of potholes in a road segment was not considered. Cracking has various levels of detail. A crack segmentation map provides the exact location of the cracks on the pavement surface, which can be used for various research and engineering purposes. For the purposes of pavement maintenance however, it is common to represent cracking simply by the total extent (in feet) and severity level of cracking (GDOT, 2007; FDOT, 2017). The extent can be for load cracking along the wheelpath, or block/transverse cracking which is not along the wheelpath. It is reasonable to assume that it is highly unlikely that smartphone sensor data can be used to accurately estimate the cracking extent regardless of crack orientation. Thus, the crack level is used to represent cracking. The crack extent and severity is calculated using the GDOT PACES protocol (GDOT, 2007). In case of multiple crack types on the same road segment, the largest crack level is used.

The raw 3D pavement data has to be processed to provide the quantified distress values. The laser scanner system includes software (LCMS RoadInspect) to calculate various pavement condition properties, including IRI, potholes, cracking, rutting and raveling using proprietary algorithms on the 3D pavement data. Additionally, existing algorithms devel-

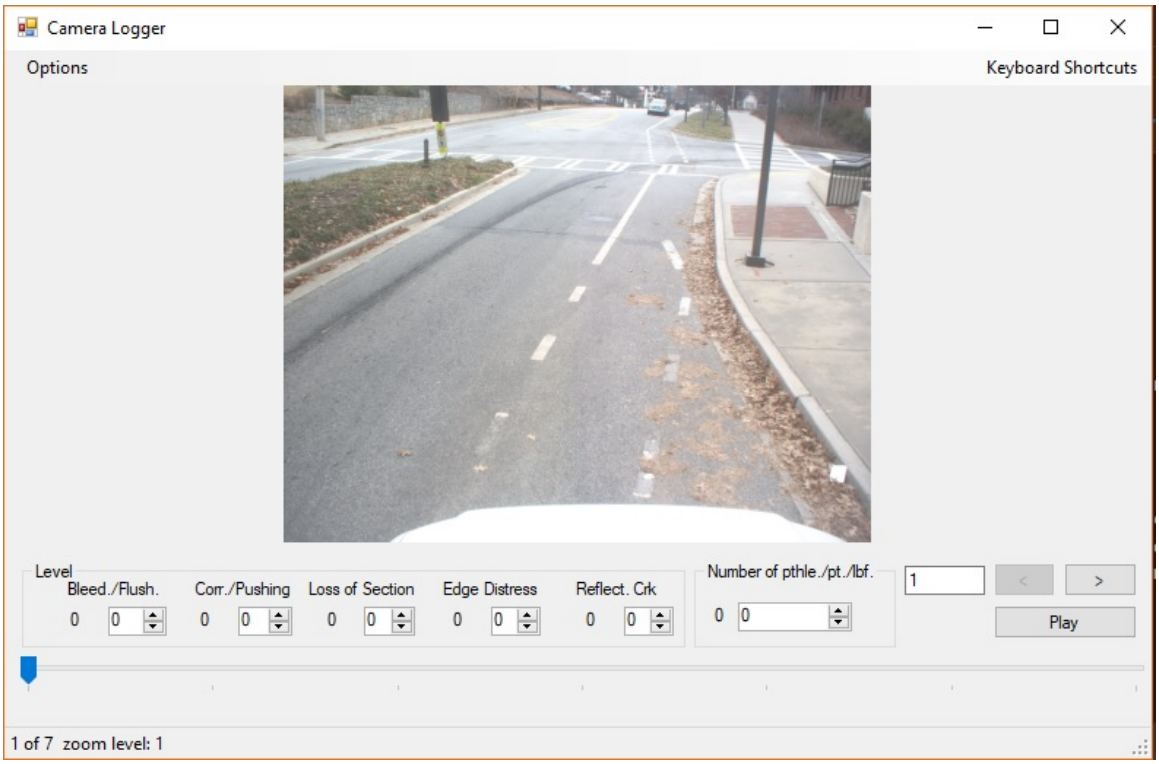



Figure 4.5: Screenshot of Camera Logger

oped by our research group can detect and classify pavement distresses such as cracking (Chatterjee, 2017; Jiang and Tsai, 2016), rutting (Tsai, Li, and Wu, 2013), raveling (Tsai and Wang, 2015) and potholes (Tsai and Chatterjee, 2017) using 3D pavement data. LCMS RoadInspect was used for estimating the IRI and raveling index. Algorithms developed by our research group were used for cracking (Chatterjee, 2017) and rutting (Tsai, Li, and Wu, 2013). Potholes were marked manually by me using the Camera Logger manual annotation tool I had developed earlier. A screenshot of the tool is given in figure 4.5. The definition of potholes/patches/local base failures definition in the GDOT PACES Manual (GDOT, 2007) was used for reference.

The IRI estimated using LCMS RoadInspect was validated by comparing it with IRI derived from certified GDOT Profilometers in the case study in section 6.1. Validation for the measurement of cracking, rutting and raveling using 3D pavement data has been completed in existing literature (Chatterjee, 2017; Tsai, Wang, and Li, 2015; Laurent, Hébert,

Lefebvre, and Savard, 2012).

### 4.1.4 Model

Existing methods for pavement condition estimation using low-cost vehicle mounted sensors (including smartphones) have often relied on supervised learning, trained using profilometer data (Dawkins, Bishop, Powell, and Bevly, 2011; Flintsch, Valeri, Katicha, Leon Izeppi, and Medina-Flintsch, 2012; Robinson, 2012). The output vector is the pavement condition metric to be determined, for example, IRI. The input vector is a feature based on the sensor data designed to provide a representation of the raw input data useful for the model to estimate the label. However, designing an input feature that encompasses all factors which affect the label is a challenging task for the given problem. A large number of factors affect the sensor response from smartphones, including vehicle model, vehicle speed, smartphone orientation and smartphone position. Although attempts have been made to incorporate factors such as vehicle speed (Zeng, 2014), several factors have not been considered by existing literature. Some examples include the effect of vertical slope on the vertical acceleration, which is generally used to estimate IRI, the impact of pavement type (Asphalt, Concrete) and the impact of presence of concrete joints, which introduce periodic bumps in the vibration response.

Deep learning concepts offer a solution to this feature selection problem by learning useful features automatically from training data. Given the large volume of labeled data provided by the laser scanners, this makes it possible to train deep models which can take the raw smartphone data as input and learn the features which affect the road condition. Deep learning has been successfully applied for the detection of pavement distresses such as cracking from vehicle-mounted cameras (Zhang, Wang, Li, Yang, Dai, Peng, Fei, Liu, Li, and Chen, 2017). However, to my knowledge, the use of deep models to estimate the pavement condition using smartphone sensors such as accelerometers has not been explored. The input data is a series of smartphone sensor measurements. Recurrent Neural

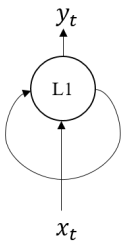



Figure 4.6: Single Layer Recurrent Neural Network

Networks (RNNs) with Long Short-Term Memory (LSTM) units have shown great performance in learning features from time series data (Williams, 2017).

*Background on Recurrent Neural Networks (RNNs) and Long Short Term Memory (LSTM) Units*

RNNs are artificial neural networks where the activations from a previous time step are fed back into the network at the next time step, allowing the network to use information from prior time steps. A simple single layer RNN is represented as a network with a loop connection as shown in figure 4.6. The loop connection represents information from the previous time step being fed back as part of the input for the current time step. Instead of this recursive representation, the RNN can be unrolled to show the inputs and outputs in each time step (figure 4.7). This unrolled representation shows how a time series input $\{x_1, x_2, ..., x_n\}$ can be fed into the network, and the information from the previous time step is passed on to the subsequent time steps. The same parameters are used for every time step. In the representation in figure 4.6, the output is obtained after a single layer. RNNs can be stacked on top of each other to allow the model to learn more complex representations (figure 4.8).

This model represents an RNN where both the input and output are series (many-to-many). The concept can be extended to convert a single input to single output (one-to-one), single input to series output (one-to-many) and series input to single output (many-to-one). RNNs have been successfully applied to several problems where the input can be modelled

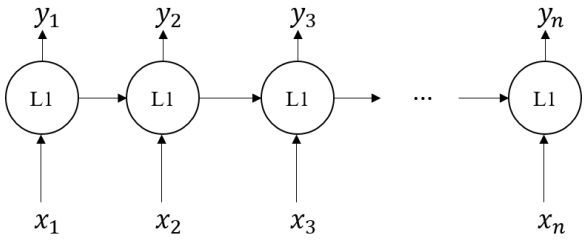



Figure 4.7: Single Layer Recurrent Neural Network (Unrolled)

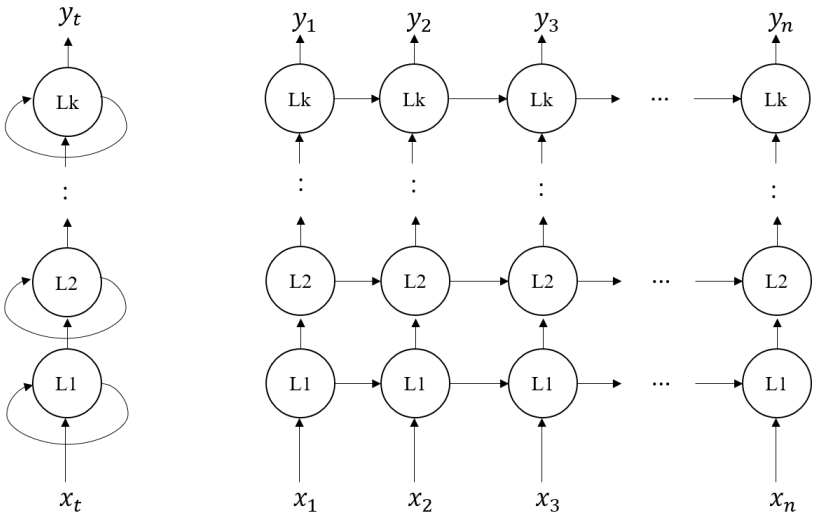



Figure 4.8: Multiple Layer Recurrent Neural Network

as a series, including speech recognition, image captioning and translation.

There are several different approaches for combining the input and information from the previous time step to calculate the output and the information for the next time step. In a simple RNN, often referred to as a "Vanilla RNN", the vector representing information from the previous time ($h_{t-1}$) step (also known as hidden state) and input are simply linearly transformed, added and passed through an activation function (4.2), which is usually the element-wise hyperbolic tangent function. The output is a linear transform of the updated hidden state (4.3).

$$h_t = tanh(W_{hh}h_{t-1} + b_{hh} + W_{xh}x_t + b_{xh}) \tag{4.2}$$

Where $W_{hh}$, $b_{hh}$, $W_{xh}$ and $b_{xh}$ are learnable parameters of the model and

$$y_t = W_{hy}h_t + b_{hy} \tag{4.3}$$

Where $W_{hy}$ and $b_{hy}$ are learnable parameters of the model.

The disadvantage of this approach is that it makes is difficult for information from previous time steps to persist after several time steps. Long Short Term Memory (LSTM) units provide a solution by providing one additional representation of information from the previous time step. This additional representation, called the cell state, is not passed through any activation function after each time step. Information is deleted and added to it only through element-wise multiplication and addition. Thus, it is possible for information added to the cell state at a time step to persist much longer.

Given an input $x_t \in \mathbb{R}^D$, the previous hidden state $h_{t-1} \in \mathbb{R}^H$ and previous cell state $c_{t-1} \in \mathbb{R}^H$, first, the activation vector $a \in \mathbb{R}^{4H}$ is computed as

$$a = W_x x_t + W_h h_{t-1} + b \tag{4.4}$$

Where $W_x$, $W_h$ and $b$ are learnable parameters. The activation vector is then split into 4 vectors $a_i, a_f, a_o, a_g \in \mathbb{R}^H$ where $a_i$ are the first $H$ elements of $a$, $a_f$ are the next $H$ elements of $a$, $a_o$ are the next $H$ elements of $a$ and $a_g$ are the last remaining elements of $a$. Next, the input gate $i$, forget gate $f$, output gate $o$ and block input gate $g$ are calculated using

$$i = \sigma(a_i) \tag{4.5}$$

$$f = \sigma(a_f) \tag{4.6}$$

$$o = \sigma(a_o) \tag{4.7}$$

$$g = \tanh(a_g) \tag{4.8}$$

Where $\sigma$ is the element-wise sigmoid function. Finally, the next cell state $c_t$ and hidden state $h_t$ are calculated using

$$c_t = f \odot c_{t-1} + i \odot g \tag{4.9}$$

$$h_t = o \odot tanh(c_t) \tag{4.10}$$

The output $y_t$ is calculated as before using equation (4.3).

*Proposed Model for Pavement Condition Estimation*

A separate network for each distress was used to make the system flexible for adjustment without affecting the other distresses. Thus, the output of the networks is a single distress value, while the input is a series of data points with each data point consisting of 12 features. A multiple layer RNN with LSTM units, followed by a fully connected layer was used, as

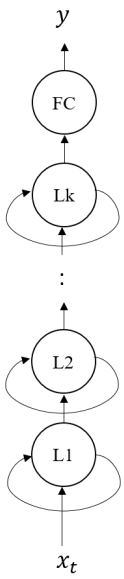



Figure 4.9: Proposed Model for Pavement Condition Estimation

shown in figure 4.9. Various values for the number of layers and size of the hidden state were tested as hyperparameters to the model for each distress. The final hyperparameter values defining the model are given in section 4.1.5.

### 4.1.5 Model Training

Before training the model parameters, the following data preprocessing steps were applied:

1. **Form input-label pairs:** The data registration methodology presented in chapter 3 was applied to assign a road segment to 3D pavement and smartphone data from each run where both 3D pavement and smartphone data were collected (table 6.6). Each 3D pavement data point and sequence of smartphone data points was assigned an index based on the run and segment indices. Thus, for each road segment in each timestamp, a corresponding input-label pair was obtained.

2. **Training-Validation-Testing Split:** The input-label pairs were split into three disjoint (but not exhaustive as explained below) subsets. The training set was used to train the model parameters. The validation set was used to assess the ability of the

model to generalize and to adjust the model hyperparameters. Finally, the performance on the test set was reported.

Generally, the dataset is randomly split into the training, validation and testing set. However, as shown in figure 3.7, two series of IRI values can have a correlation more than the noise even for shifts of up to $k = 10$, which corresponds to 50 meters of shift on either side. This indicates that if a particular road segment is in the training set, then there should be no validation or testing data on either side of it for up to 50 meters. Therefore, the data collection route was split into training and validation/testing bands with buffers between them, as shown in figure 4.10. The relative sizes of the training, validation and test bands were a typical ratio of $8 : 1 : 1$.

3. **Normalization:** The input features should be normalized to avoid providing more weightage to features with larger variance. Each feature in all inputs were normalized by subtracting its mean and dividing by its standard deviation.

$$x_d^{(i)} = \frac{x_d^{(i)} - \bar{x_d}}{\sigma_{x_d}} \tag{4.11}$$

Where $\bar{x_d}$ and $\sigma_{x_d}$ are the mean and standard deviation of the values of feature $d$ respectively. $\bar{x_d}$ and $\sigma_{x_d}$ of each feature was calculated using the inputs in the training set only.

4. **Shuffle Data:** The input-label pairs were shuffled to ensure that mini-batches used for training the model represent diverse cases of pavement condition.

The bulk of the computation effort for training was carried out using GPU acceleration. An NVIDIA 1080Ti 16GB GPU was used for training the models. The parameters of the proposed models for each distress were trained using backpropagation. A stochastic gradient descent optimizer was used for updating the weights, with a weight decay of 0.1 for regularization. Minibatches of size 500 were used for training. The number of RNN

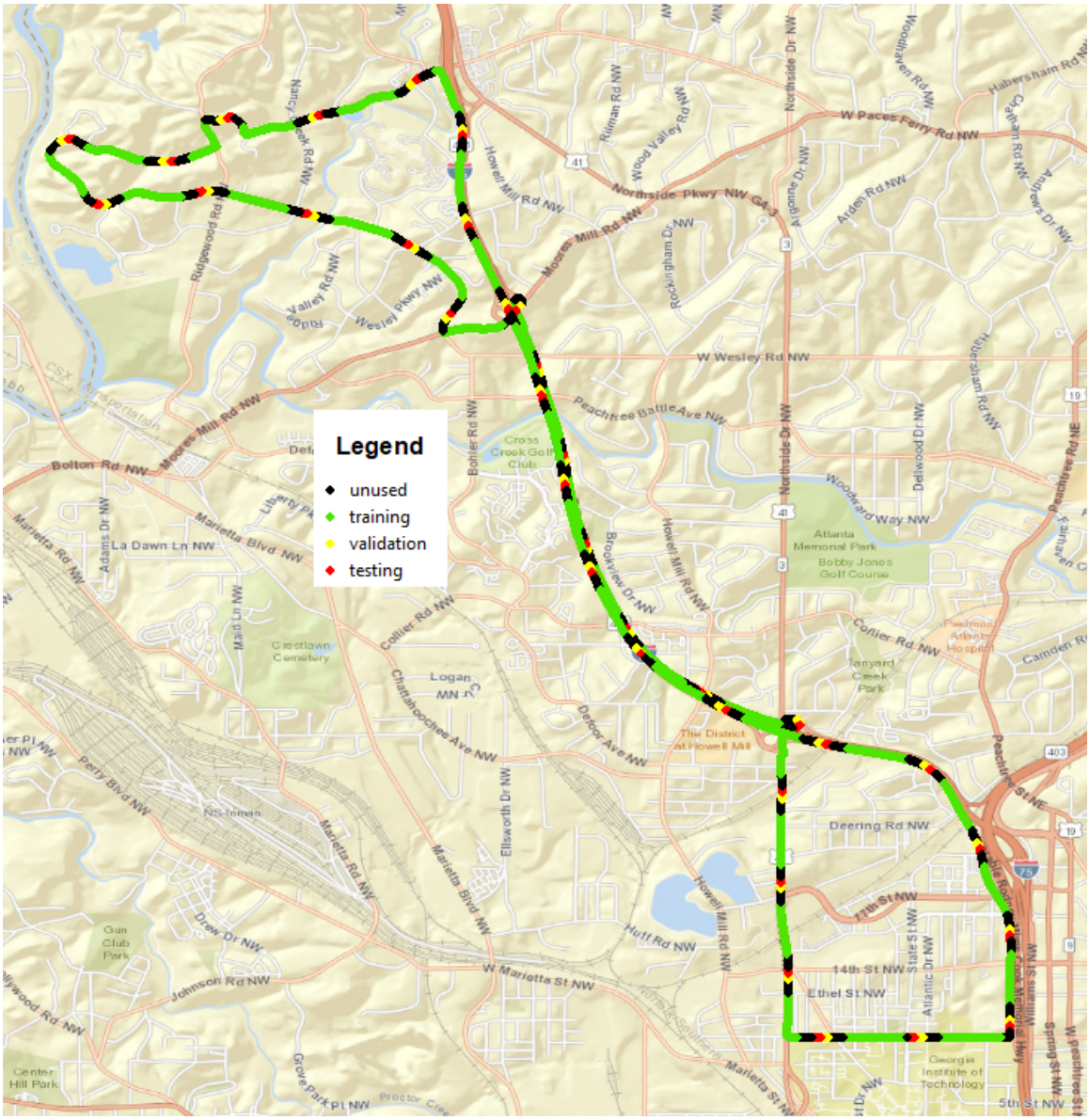



Figure 4.10: Training-validation-testing split of the data collection route

layers, learning rate schedule and number of epochs were specific to the model for each distress, which are explained in detail in the subsections below. For each distress, the epoch with the lowest validation error was saved as the model.

*Model Training for IRI*

The model for IRI estimation consisted of 10 layer RNN with LSTM units. The number of features in the hidden state and cell state was 16 for each layer. The learning rate was set to 0.1 for the first 90 epochs, 0.01 for the next 90 epochs and 0.001 for the last 90 epochs for a total of 270 epochs. The training and validation loss curves are shown in figure 4.11.

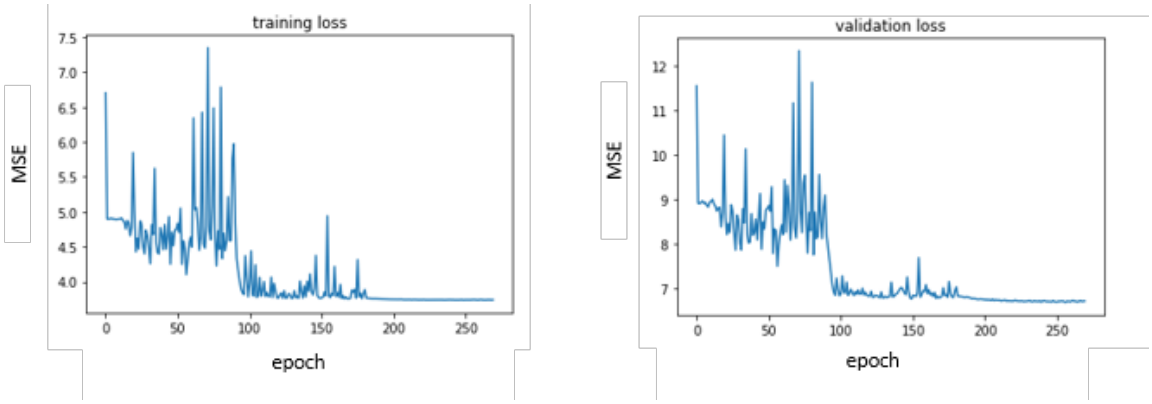



Figure 4.11: Training and validation loss curves for IRI Model

*Model Training for Rutting*

The model for rutting estimation consisted of a 2-layer RNN with LSTM units. The number of features in the hidden state and cell state was 5 for each layer. Increasing the complexity of the model did not produce any significant improvement in performance. The learning rate was set to 0.1 for the first 50 epochs, 0.01 for the next 50 epochs and 0.001 for the next 50 epochs, for a total of 150 epochs. The training and validation loss curves are shown in figure 4.12.

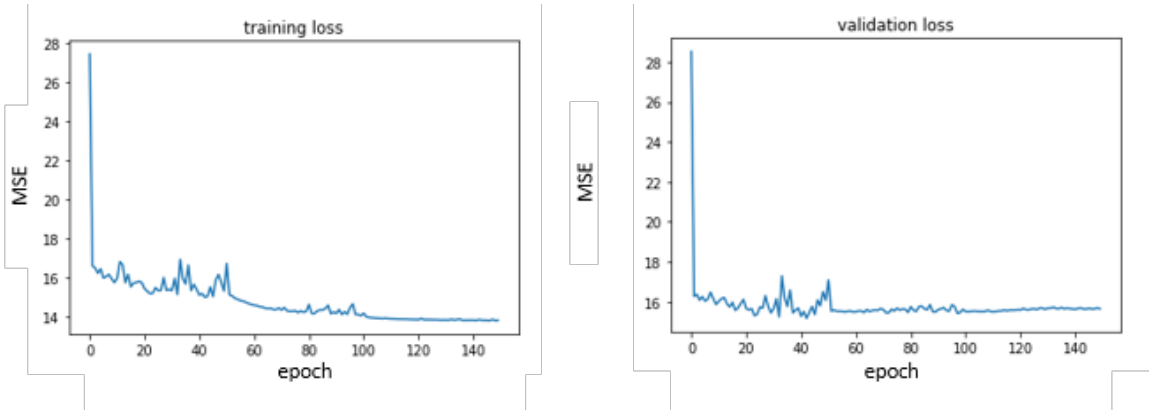



Figure 4.12: Training and validation loss curves for rutting model

*Model Training for Raveling*

The model for raveling estimation consisted of a 2-layer RNN with LSTM units. The number of features in the hidden state and cell state was 8 for each layer. The learning rate was set to, 0.001 for the first 50 epochs, 0.0001 for the next 100 epochs and 0.00001 for the next 200 epochs, for a total of 350 epochs. The training and validation loss curves are shown in figure 4.13.

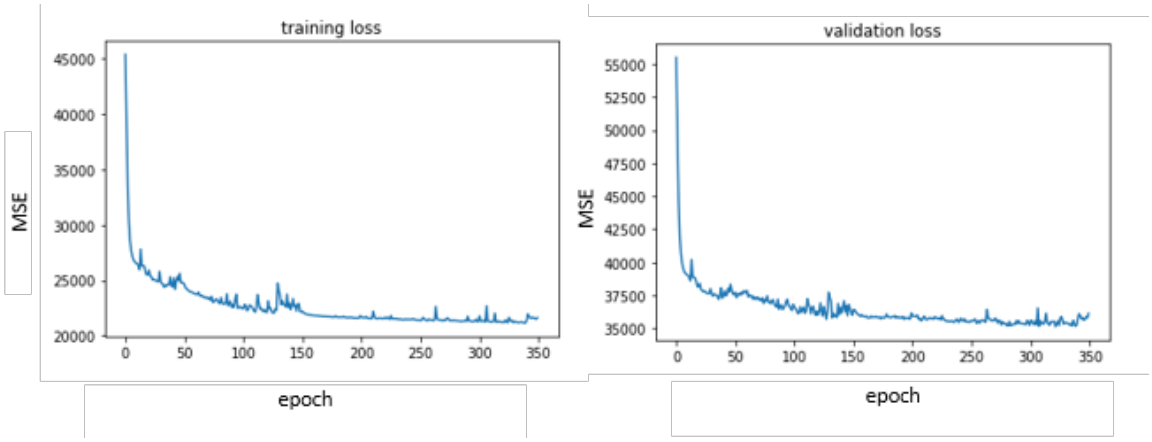



Figure 4.13: Training and validation loss curves for raveling model

*Model Training for Cracking*

The model for cracking estimation consisted of a 4-layer RNN with LSTM units. The number of features in the hidden state and cell state was 8 for each layer. The learning rate

was set to, 0.01 for the first 90 epochs, 0.001 for the next 270 epochs and 0.0001 for the last 180 epochs, for a total of 540 epochs. The weight decay was lowered to 0.0001. The training and validation loss curves are shown in figure 4.14.

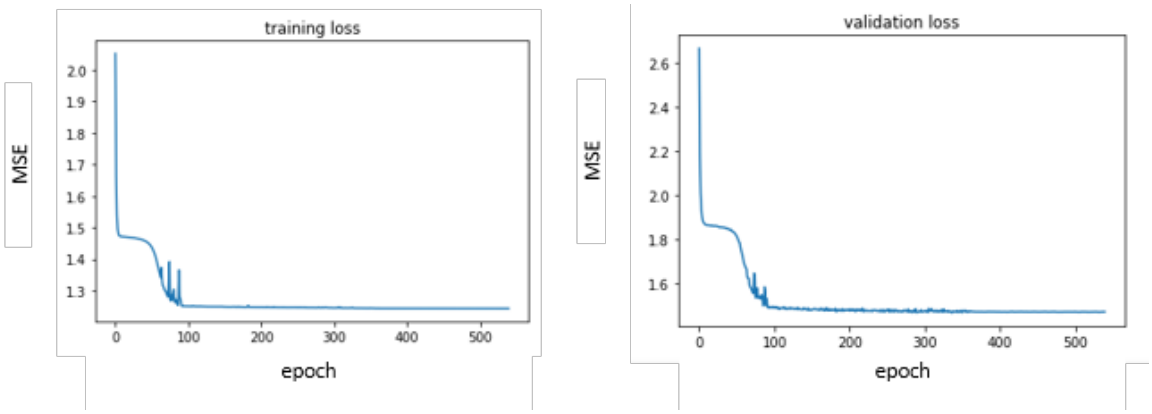



Figure 4.14: Training and validation loss curves for cracking model

*Model Training for Potholes*

The model for pothole estimation consisted of a 2-layer RNN with LSTM units. The number of features in the hidden state and cell state was 8 for each layer. The learning rate was set to 0.001 for 1000 epochs. The training and validation loss curves are shown in figure 4.15.

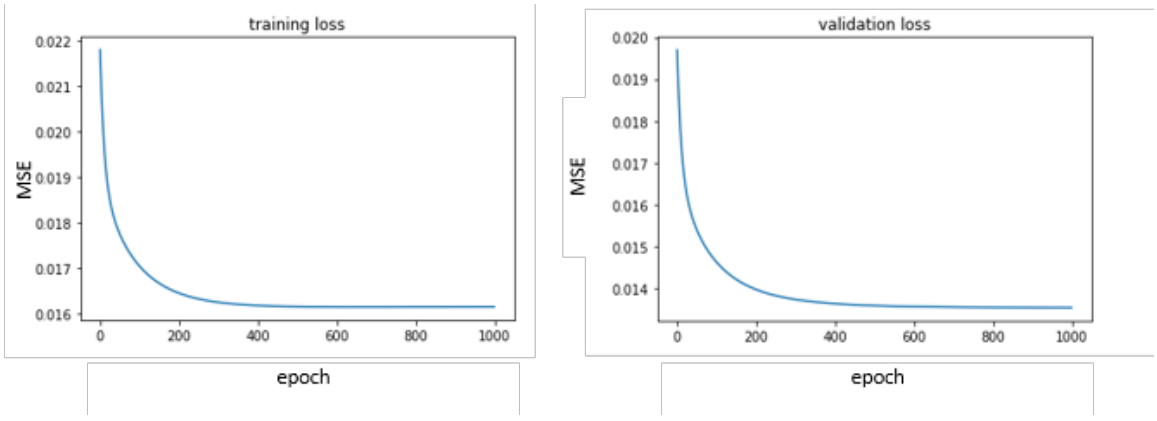



Figure 4.15: Training and validation loss curves for pothole model

Table 4.2: Root mean squared error on test set

| Distress Value | RMSE |
|---|---|
| IRI ($m/km$) | 1.7483 |
| Average Rut Depth ($mm$) | 4.1313 |
| Raveling Index ($cm^3/m^2$) | 131.3459 |
| Maximum Crack Level | 1.2104 |
| Pothole Probability | 0.1269 |

## 4.2 Results and Validation of Proposed Methodology

The overall root mean squared error (RMSE) on the test set by each model is shown in table 4.2. However, the RMSE can be heavily skewed due to the presence of outliers with large errors, as shown by the distribution of the test errors for IRI in figure 4.16 for 1 run. The outliers are more clearly visible in the boxplot in figure 4.17. Figure 4.18 demonstrates how removing a small percentage of the bad pavement condition estimates can drastically improve the performance of the model. For example, removing just the 10% worst results reduces the IRI RMSE by 48.6% to 0.8991 m/km. This underscores the importance of separating good and bad pavement condition estimates, which is explored in chapter 5. There are other factors, such as smoothing (section 6.6), which have to be considered as well when interpreting the performance of the proposed methodology. After considering these factors, the feasibility of the methodology proposed in this chapter for use in the field is discussed in section 6.7.

As shown in chapter 2, there are three major shortcomings in existing methods of pavement condition evaluation using smartphones: 1) the existing methods are highly sensitive to the user context, such as vehicle model, vehicle speed, smartphone orientation and position, etc.; 2) the existing methods focus mainly on just two distresses: roughness and potholes; and 3) there is no confidence level associated with the pavement condition estimations which diminishes their practical utility. The first and second shortcomings are addressed in this chapter, while the third shortcoming is addressed in chapter 5 to drastically improve pavement condition estimation using crowdsourced smartphone data.

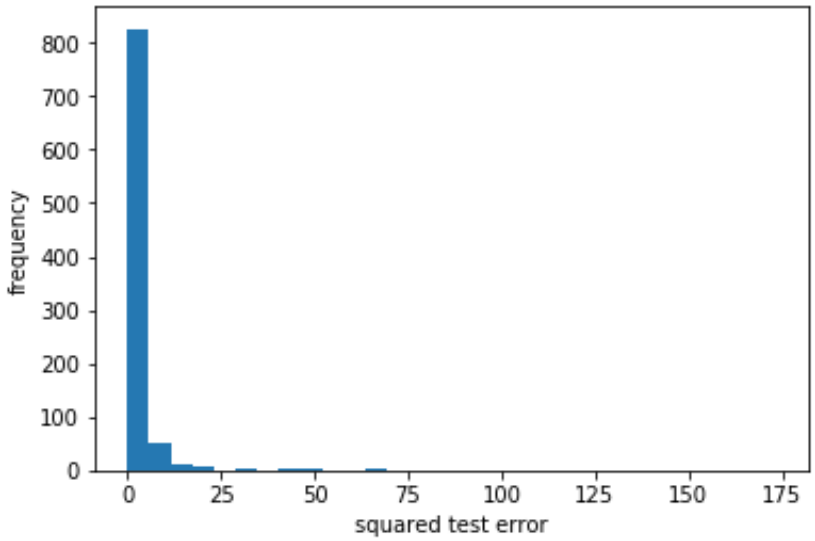



Figure 4.16: Distribution of IRI error in test set, highlighting outliers

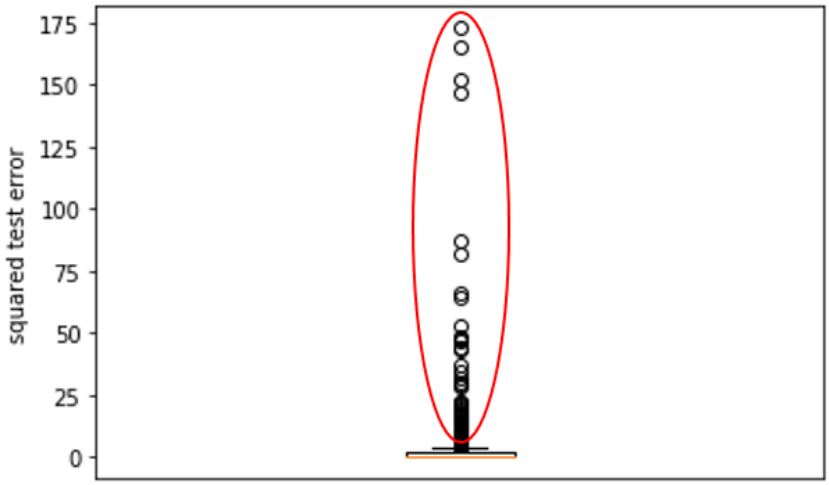



Figure 4.17: Box plot of IRI error in test set

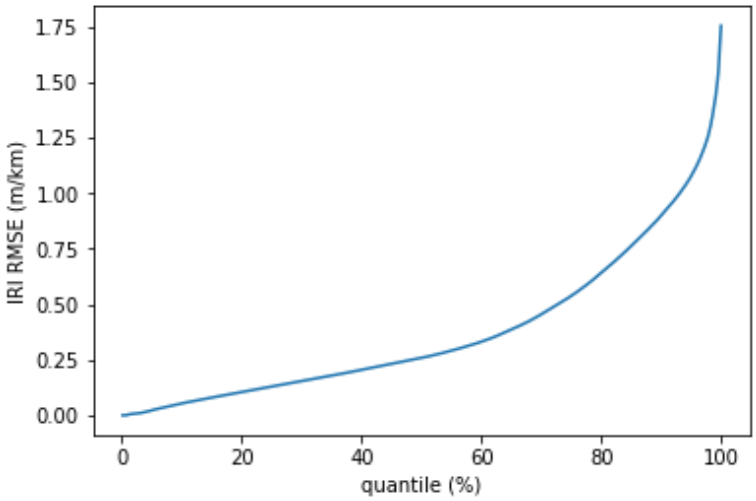
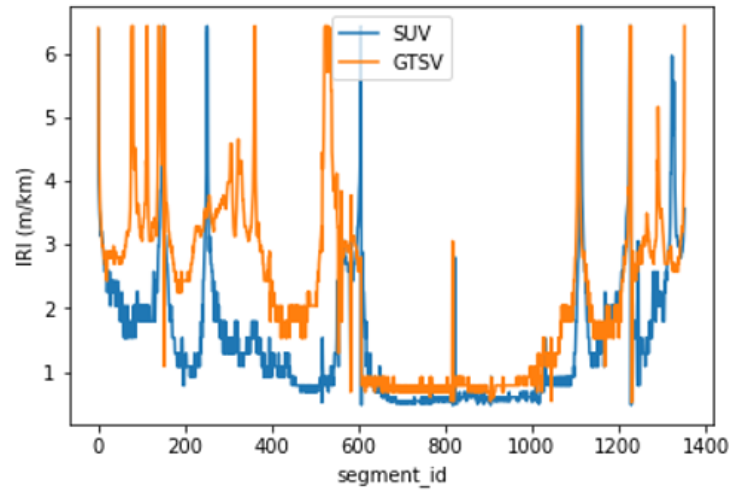


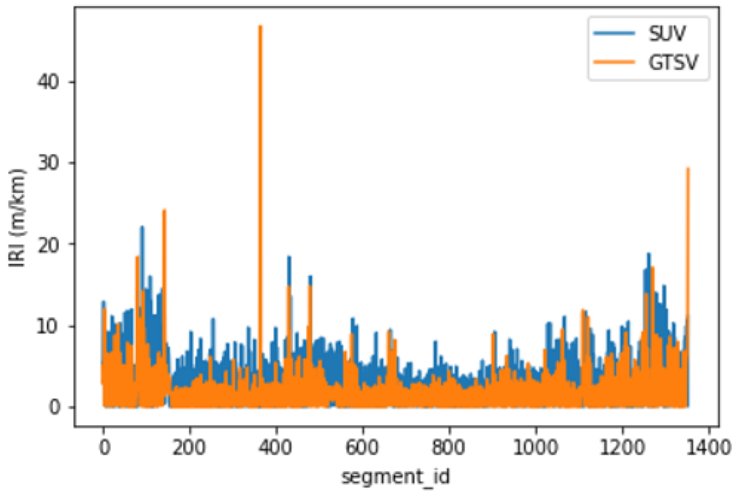

Figure 4.18: IRI RMSE after taking removing errors larger than the quantile value in the x-axis

To assess the sensitivity of the proposed deep learning based model for pavement condition estimation, the IRI was estimated for the same route using the proposed methodology and using a pseudo-IRI model widely used in existing literature on two different vehicles: the GTSV and an SUV. The repeatability of the estimated IRI from each model was then compared. A smaller difference between the pavement condition estimated from the two vehicles is preferred. Figure 4.19 demonstrates the estimated IRI along the route. Figure 4.19a compares the IRI calculated from data collected from each vehicle using the proposed model. Figure 4.19b shows the same for the pseudo-IRI model. Figure 4.19c shows the pseudo-IRI model results after smoothing with a moving average filter of window size 250 m for clarity. The RMSE between the two pavement condition estimates from the proposed model is 1.38 $m/km$, which is significantly lower than that from the pseudo-IRI model (4.05 $m/km$), thus indicating that the proposed model is less sensitive to factors outside the pavement condition such as vehicle model as compared to the popular existing approach, which is desirable.

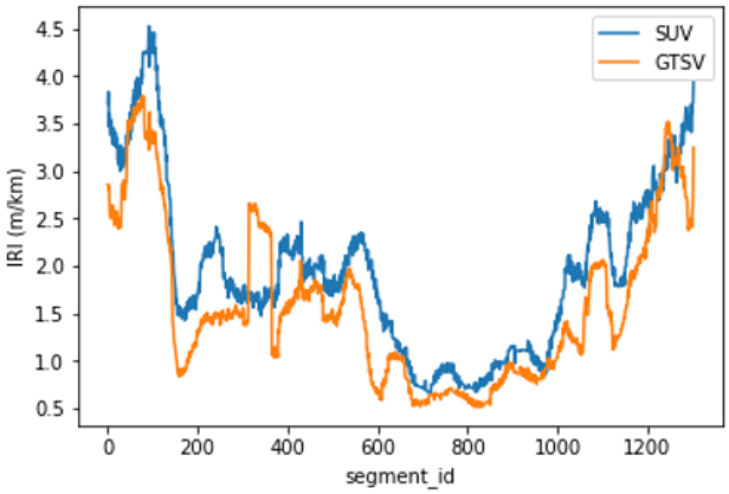
This chapter presented and validated a methodology for single-run pavement condition estimation. In chapter 5, the methodology for combining pavement condition estimates




(a)




(b)




(c)

Figure 4.19: Estimated IRI from (a) proposed model, (b) pseudo-IRI model and (c) pseudo-IRI model (smoothed)

from multiple runs of crowdsourced smartphone data is presented.

# CHAPTER 5
# METHODOLOGY FOR MULTIPLE RUN PAVEMENT CONDITION ESTIMATION AND PREDICTION USING SMARTPHONE AND 3D PAVEMENT DATA

In chapter 4, the methodology for estimating the pavement condition using smartphone data collected in a single run was presented. In this chapter, the methodology for estimating the pavement condition using smartphone and 3D pavement data from multiple runs is presented. First, a methodology for combining single-run pavement condition estimates from a single time step to produce a multiple-run pavement condition estimate with confidence level is presented. Second, a methodology for enhancing pavement condition forecasting using prior 3D pavement and smartphone data combined with crowdsourced smartphone data is presented.

## 5.1 Proposed Methodology for Multiple Run Pavement Condition Estimation with Confidence Level

In a given time step, pavement condition data for a given road segment can be collected from multiple runs. Using the methodology presented in chapter 4, the pavement condition can be estimated from each of these runs. However, each run will probably provide a different estimate of the pavement condition, including some outliers which provide highly inaccurate pavement condition estimates. A method is needed to combine these single-run pavement condition estimates and also associate a confidence level with each estimate which help differentiate between good and bad pavement condition estimates.

The proposed methodology for multiple-run pavement condition estimation with confidence level is illustrated in figure 5.1. First, the collected data is registered to the GIS model of the road network using the methodology presented in chapter 3. Second, the

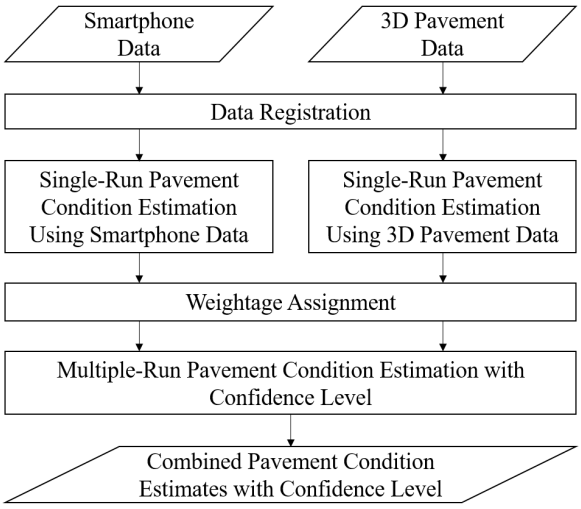



Figure 5.1: Flowchart of proposed methodology for multiple-run pavement condition estimation

pavement condition is estimated from each run individually. The algorithms described in section 4.1.3 are used to estimate the pavement condition from a single run of 3D pavement data. The trained models for pavement condition estimation using smartphone data (chapter 4) are used to estimate the pavement condition using smartphone data.

Third, the method for estimating the pavement condition using multiple-run data should be generalized to use data from multiple types of sensors, including smartphones and 3D pavement laser scanners, with the appropriate weightage given to each type of sensor. The pavement condition estimated by the 3D pavement data is expected to be more accurate than that provided by smartphone sensors and thus, should be given a higher weightage. A weight of 1 is assigned to the pavement condition estimates from 3D pavement data and a lower weight $\epsilon < 1$ is assigned to pavement condition estimates from smartphone data. $\epsilon$ is a hyperparameter which will be tuned with the help of the case studies detailed in chapter 6.

Finally, a correlation averaging based method was implemented based on previous work by Ndoye, Barker, Krogmeier, and Bullock (2011) for combining the single-run pavement condition estimates. The following improvements to the original correlation averaging method were implemented:

1. In this study, the data points are distributed to multiple road segments during data registration. For a given road segment, only the data points corresponding to the given road segment were combined from multiple runs.

2. Higher weightage was given to pavement condition estimates derived from 3D pavement data as described above. This concept can be generalized to accommodate more sources of pavement condition data with different weights.

3. A confidence level was assigned to the combined pavement condition estimate, as explained below.

Generally, given the distribution of a test statistic $\bar{Y}$ estimating a parameter $\theta$ without bias such that $P(\bar{Y}) \sim N(\theta, \sigma^2)$, the $\gamma\%$ confidence interval can be calculated as $[\theta - z\sigma, \theta + z\sigma]$ where $z$ is the standard normal distribution critical value of $\left(1 - \frac{\gamma}{100}\right)/2$. In the case of pavement condition estimation however, from an engineering standpoint, it is the margin of error for different distresses that is better defined (GDOT, 2007; FDOT, 2017). Thus, given the sample variance which can be used to estimate the population variance without bias, the confidence level for this margin of error can be determined.

Let the acceptable margin of error for a distress value be $t$. For a road segment with unknown pavement condition distress value $\theta$, the acceptable interval becomes $[\theta - t, \theta + t]$. Thus, solving for confidence level $\gamma$,

$$z\sigma = t \tag{5.1}$$

$$\Rightarrow z = \frac{t}{\sigma} \tag{5.2}$$

Where $z$ is the standard normal distribution critical value of $\left(1 - \frac{\gamma}{100}\right)/2$. Hence,

$$\Rightarrow \frac{1 - \frac{\gamma}{100}}{2} = F\left(\frac{-t}{\sigma}\right) \tag{5.3}$$

Where $F$ is the CDF of the standard normal distribution. Thus, the confidence level $\gamma\%$ of the combined pavement condition estimate is a function of the variance $\sigma^2$ of the estimates.

$$\gamma = 100(1 - 2F(-t/\sigma)) \tag{5.4}$$

Note that in our case, each single-run pavement condition estimate is weighted. Therefore, the weighted variance of the sample must be used. For a given road segment, if multiple single-run distress value estimates $y_i \forall i \in \{1, ...n\}$ with different weights $w_i$ are collected forming a sample, the weighted variance of this sample is given by

$$\sigma_s^2 = \left( \frac{\sum_{i=1}^{n} \left[ w_i \left( y_i - \bar{y} \right)^2 \right]}{\sum_{i=1}^{n} w_i} \right) \tag{5.5}$$

where

$$\bar{y} = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i} \tag{5.6}$$

Thus, using the proposed methodology, the pavement condition for given road segments can be estimated using multiple-run smartphone and 3D pavement data with an associated confidence level.

## 5.2 Proposed Methodology for Pavement Condition Prediction

### 5.2.1 Background

Pavement condition assessment using 3D pavement data is typically carried out only once a year or once every two years. However, especially for rapidly deteriorating road segments, more frequent (monthly or quarterly) pavement condition assessment is desired. Haider, Baladi, Chatti, and Dean (2010) concluded that more frequent pavement condition assessment was needed for pavement functional performance measurement, and that a higher frequency of data collection will affect the MR&R decision making.

In section 5.1, it was assumed that the observed data was collected over a short duration (one time step) where it was reasonable to assume that there was negligible change in the pavement condition, for e.g. within a day or within a week. One approach for continually evaluating the pavement condition over multiple time steps is to simply group the collected data into discrete time steps and independently estimate the pavement condition at each time step.

However, this frequentist approach has several shortcomings:

1. To estimate the pavement condition for a specific time interval, all data collected in the previous time intervals for the same road segments is not leveraged. This shortcoming is especially a problem when we consider the use case described above, where smartphone data is applied to estimate the pavement condition between low frequency 3D pavement data collection, which means information from accurate 3D pavement data is discarded when estimating the pavement condition for the same road segments in subsequent time steps.

2. The frequentist approach is prone to outliers, especially in time intervals of road segments where a small number of runs were collected.

Using a Bayesian approach overcomes these shortcomings by leveraging the data collected in the past for the analyzed road segments. In a Bayesian approach, prior information about the road segment is also used as an input along with the collected data to predict the pavement condition, solving the first shortcoming. A Bayesian approach is also more robust against small sample sizes in the observed data generally by giving the observed data a weightage based on the observed data sample size. Thus, for small sample sizes which are more prone to outliers, more weightage will be given to the prior information, mitigating the second shortcoming.

This methodology is described for a generic "distress value": a scalar quantity describing the pavement condition, such as IRI, average rut depth or raveling index. Thus, it can

be applied to any scalar measure of pavement condition which can potentially be captured by vehicle vibration. In section 5.3, the proposed methodology is validated by applying it to IRI estimation.

Suppose at some time step, we know the probability distribution of a distress value for a given road segment, defined by a random variable $\theta$. Let $X$ be the random variable defining the distress value we will observe at this time step for that road segment. Then our knowledge about the probability distribution of $\theta$ can be updated using the evidence $X$ and prior $\theta$ using Bayes theorem:

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{\int P(X|\theta)P(\theta)d\theta} \tag{5.7}$$

Where $P(\theta|X)$ is the posterior marginal probability given the evidence (the updated probability distribution of the distress value after considering the evidence), $P(X|\theta)$ is the likelihood of observing the evidence given the prior and $P(\theta)$ is the prior probability. To use Bayes theorem to update our beliefs of the distress value, it has to be modeled as a random variable. However, in practice, a deterministic value is preferred. The expectation of the distribution of the distress value after considering the evidence ($E[\theta|X]$) can be reported as the final predicted distress value for a time step.

$$E[\theta|X] = \int_{-\infty}^{\infty} \theta P(\theta|X)d\theta \tag{5.8}$$

The flowchart in figure 5.2 describes the steps for iteratively predicting the pavement condition given the prior pavement condition and newly collected evidence at each time step. Each step is explained in the subsections below.

### 5.2.2 Initial prior pavement condition estimation

First, the initial prior probability of the distress value needs to be determined. As shown in section 5.2.4, the prior and posterior probabilities will follow a normal distribution. In

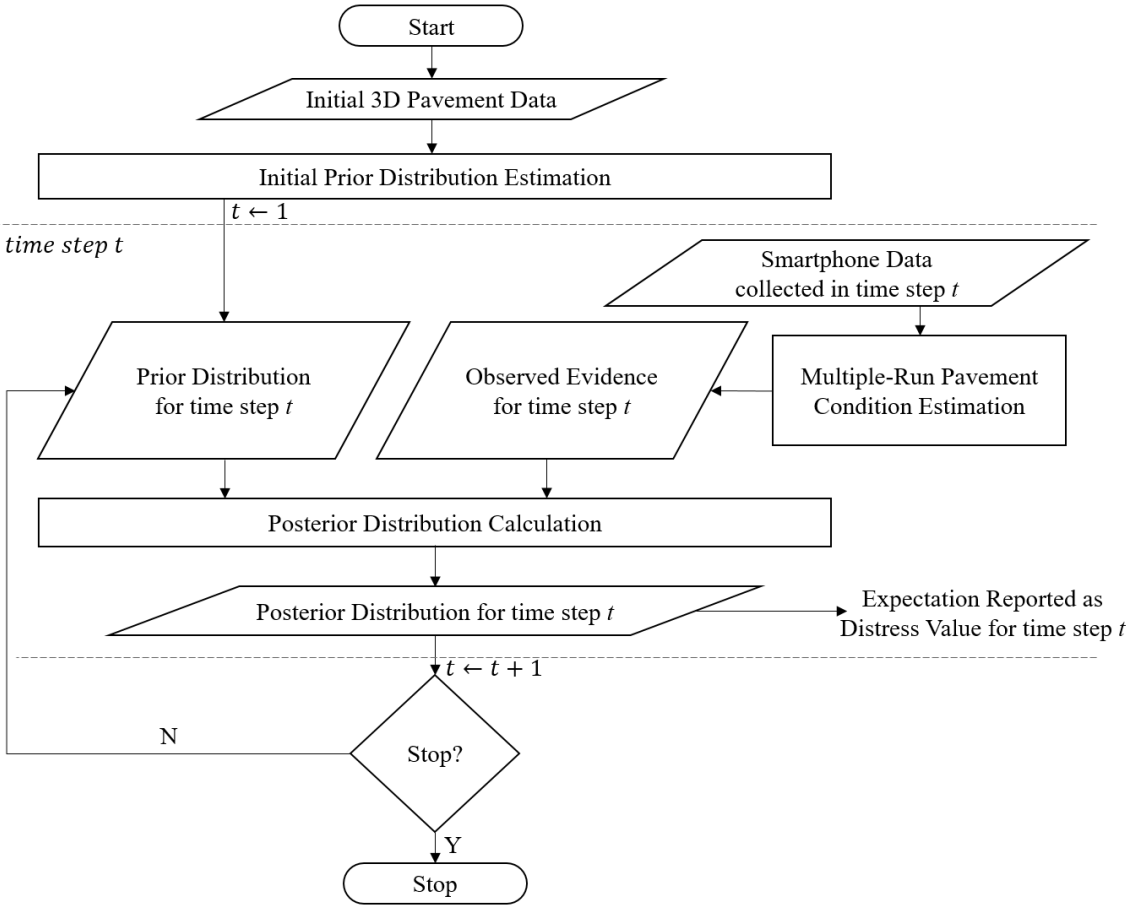



Figure 5.2: Flowchart of proposed methodology for pavement condition prediction

our use case (section 5.2.1), the initial distress value is expected to be provided by a single highly accurate estimate from 3D pavement data $m_0$. As explained in section 5.1, the single-run pavement condition estimates are given a weightage based on the sensor that provided the data for the estimate. That weightage will be useful in assigning an effective sample size for the prior $n_0$ (Lenth, 2001). For 3D pavement data, the weight was 1, giving an effective sample size of $1 \times 1 = 1$. In section 5.2.4, it is explained how the effective sample size is used to calculate the variance of the initial prior distribution $s_0^2$.

$$s_0^2 = \frac{1}{n_0} = 1 \tag{5.9}$$

Thus, the initial prior distress value distribution is given by a normal distribution with mean $m_0$ and variance of 1. The hyperparameter $\epsilon$ (section 5.1), which refers to the weight of the distress value estimated from a single run of smartphone data, will control the weightage given to the initial distress value estimate. This hyperparameter will be tuned to minimize the error in pavement condition prediction in section 5.3.

### 5.2.3 Multiple-Run Pavement Condition Estimation

Second, for each time step $t$, smartphone data is collected for the concerned road segments from multiple runs. The methodology presented in section 5.1 is used to estimate the distress value at time step $t$ without the use of prior information. In our use case, equally weighted distress value estimates from registered smartphone data $x_1, x_2, ..., x_n$ are combined to give the observed distress value estimate $\bar{x}$ for time $t$.

$$\bar{x} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} = \frac{\epsilon \sum_{i=1}^{n} x_i}{n\epsilon} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{5.10}$$

Suppose $n$ runs are collected for a given road segment. Each run of smartphone data is given a weightage of $\epsilon$, as explained in section 5.1, giving an effective sample size of $n \times \epsilon = n_x$.

### 5.2.4 Posterior Pavement Condition Estimation

Given the likelihood probability distribution, it is preferable to have a conjugate prior probability distribution. Conjugate prior distributions have the desirable property of producing a posterior probability distribution of the same family of distributions as the prior probability distribution. Thus, once derived, the equation for calculating the posterior distribution can be iteratively used in each time step to update the probability distribution of the distress value. The multiple-run pavement condition estimate is obtained as a weighted average of the single-run pavement condition estimates, which are independent experiments. Thus, by the Central Limit Theorem, the likelihood probability distribution in our case follows a normal distribution. There are two suitable conjugate prior distributions that can be considered. Both models were implemented and analyzed in this thesis.

1. Normal distribution with known likelihood variance.

2. Normal distribution with unknown likelihood variance.

In both cases, the posterior probability of the distress value is calculated to be a normal distribution. The expectation of this distribution is returned as the predicted pavement condition for the current time step with the variance of the distribution used to calculate the confidence level of the prediction as explained in section 5.1. If this is the last time step, the algorithm exits. Otherwise, the posterior distribution is used as the prior distribution for the next time step and the process is repeated with new data collected in the next time step. The subsections below explain each model in detail.

*Normal prior distribution with known likelihood variance*

In this model, the likelihood follows a normal distribution with known variance $\sigma_x^2$.

$$p(X|\mu) \sim N(\mu, \sigma_x^2) \tag{5.11}$$

where $\mu$ is a random variable and $\sigma_x^2$ is a constant.

Then the conjugate prior distribution for $\mu$ is a normal distribution with known parameters $m$ and $s^2$.

$$P(\mu) \sim N(m, s^2) \tag{5.12}$$

Applying Bayes' theorem and solving for the posterior gives another normal distribution.

$$P(\mu|X) \sim N\left(\frac{\frac{n_x\bar{x}}{\sigma_x^2} + \frac{m}{s^2}}{\frac{n_x}{\sigma_x^2} + \frac{1}{s^2}}, \frac{1}{\frac{n_x}{\sigma_x^2} + \frac{1}{s^2}}\right) \tag{5.13}$$

The formal algorithm for pavement condition prediction for a single road segment when the likelihood variance is known is given in algorithm 2.

**Algorithm 2** Pavement condition prediction with known likelihood variance

Let $\sigma_x^2$ be the known variance of the likelihood distribution
Let $m_0$ be the initial distress value estimate using 3D pavement data
$m \leftarrow m_0$
$s^2 \leftarrow 1$
**for** $t \in \{1, ..., T\}$ **do**
    Let $x_1, x_2, ..., x_n$ be the single-run distress value estimates from collected smartphone data in time step $t$.
    $\bar{x} \leftarrow (\sum_{i=1}^{n} x_i)/n$
    $n_x \leftarrow n\epsilon$
    $m \leftarrow \left(\frac{n_x\bar{x}}{\sigma_x^2} + \frac{m}{s^2}\right) / \left(\frac{n_x}{\sigma_x^2} + \frac{1}{s^2}\right)$
    $s^2 \leftarrow \left(\frac{n_x}{\sigma_x^2} + \frac{1}{s^2}\right)^{-1}$
    $d_t \leftarrow m$
**return** $d_1, d_2, ..., d_T$ ▷ return distress values for each time step

*Normal distribution prior with unknown variance*

In this model, the likelihood follows a normal distribution with unknown variance.

$$p(X|\mu, \sigma^2) \sim N(\mu, \sigma^2) \tag{5.14}$$

where $\mu$ and $\sigma^2$ are random variables.

Then the conjugate prior distribution is

$$P(\mu) \sim N(m, \sigma^2/w) \tag{5.15}$$

$$P(\sigma^2) \sim \Gamma^{-1}(\alpha) \tag{5.16}$$

where $m, w$ and $\alpha$ are known parameters and $\Gamma^{-1}$ is the inverse gamma distribution. The probability density function of the inverse gamma distribution with parameter $\alpha$ is

$$p(a) = \frac{1}{\Gamma(\alpha)} \left(\frac{1}{a}\right)^{\alpha+1} e^{-1/a} \tag{5.17}$$

where $\Gamma$ is the gamma function.

$$\Gamma(a) = \int_0^\infty k^a e^{-k} dk \tag{5.18}$$

Applying Bayes' theorem and solving for the posterior gives the following posterior distributions for $\mu$ and $\sigma^2$.

$$P(\sigma^2|X) \sim \Gamma^{-1}\left(\alpha + \frac{n_x}{2}\right) \tag{5.19}$$

$$P(\mu|\sigma^2, X) \sim N\left(\frac{n_x\bar{x} + wm}{n_x + w}, \frac{\sigma^2}{n_x + w}\right) \tag{5.20}$$

In this formulation, the parameter $w$ acts as the effective sample size of the prior. So the initial value for the parameter $w$ is 1. The initial values for parameter $\alpha$ will be experimentally determined in section 5.3. The formal algorithm for pavement condition prediction for a single road segment when the likelihood variance is unknown is given in algorithm 3.

**Algorithm 3** Pavement condition prediction with unknown likelihood variance

Let $m_0$ be the initial distress value estimate using 3D pavement data
Let $\alpha_0$ be known initial parameter
$m \leftarrow m_0$
$w \leftarrow 1$
$\alpha \leftarrow \alpha_0$
**for** $t \in \{1, ..., T\}$ **do**
    Let $x_1, x_2, ..., x_n$ be the single-run distress value estimates from collected smartphone data in time step $t$.
    $\bar{x} \leftarrow (\sum_{i=1}^{n} x_i)/n$
    $n_x \leftarrow n\epsilon$
    $\alpha \leftarrow \alpha + (n_x/2)$
    $m \leftarrow \frac{n_x\bar{x}+wm}{n_x+w}$
    $w \leftarrow w + n_x$
    $d_t \leftarrow m$
**return** $d_1, d_2, ..., d_T$ ▷ return distress values for each time step

## 5.3 Results and Validation of Proposed Methodology

### 5.3.1 Results of proposed methodology for multiple-run pavement condition estimation

To assess the performance of the proposed methodology for multiple-run pavement condition estimation, 3D pavement data was collected for the GT Test Route, along with multiple laps of the GT Test Route collecting smartphone data only to simulate crowdsourcing of pavement condition data. Only the smartphone data was used for pavement condition estimation. The 3D pavement data was used to establish the ground truth pavement condition to be estimated by the smartphone data. IRI was estimated for each single run using the methodology presented in chapter 4 and the combined to give a multiple-run pavement condition estimate with confidence level following the methodology presented in section 5.1.

The IRI estimated from individual runs and the combined IRI is given in figure 5.3. Individual runs often contain outliers, which are mitigated in the multiple run pavement condition estimate. For example, some sudden spikes are observed in run 4, which do not correspond to any feature in the ground truth (figure 5.4). In both cases, the spike appears to

be because of large vibrations captured while changing lanes for exits/intersections. Clearly these are undesirable spurious measurements from a single-run. In the combined pavement condition estimate, the effect of these outliers is mitigated. The RMSE of that individual run from the ground truth IRI is 2.3281. The RMSE of the combined pavement condition estimate is 2.2667.

Figure 5.5 shows the combined estimated IRI as compared to the ground truth IRI along the GT Test Route. As observed in figure 3.5, the high frequency changes in the ground truth is not noise, but actual fluctuations in the IRI value. It appears that the proposed method fails to capture these high frequency fluctuations although it appears to capture the low frequency trend of IRI value. The proposed methodology estimates the road condition for each road segment independently, although the pavement condition of adjacent road segments themselves can be correlated, as observed in figure 3.7.

### 5.3.2 Results of proposed methodology for confidence level assignment

As observed in figure 5.5, the proposed methodology gives a mix of good and bad estimates of the pavement condition. Even a small number of bad estimates can heavily skew the overall RMSE to a high value. It may not always be possible to improve the accuracy of pavement condition estimation using smartphone data, as the 3D pavement data collected by laser scanners carries much more information about the pavement surface which the vehicle vibration cannot capture.

One solution to this problem is to isolate and reject the bad pavement condition estimates from smartphone data, drastically reducing the RMSE of the results and making them more useful for engineers. Rejecting some of the estimates will mean that no pavement condition is reported for some road segments for a given time step, which is acceptable if crowdsourced pavement condition estimation is seen as a method to reduce manual inspection effort, not completely replace it.

To separate the good and bad pavement condition estimates, a confidence level was as-

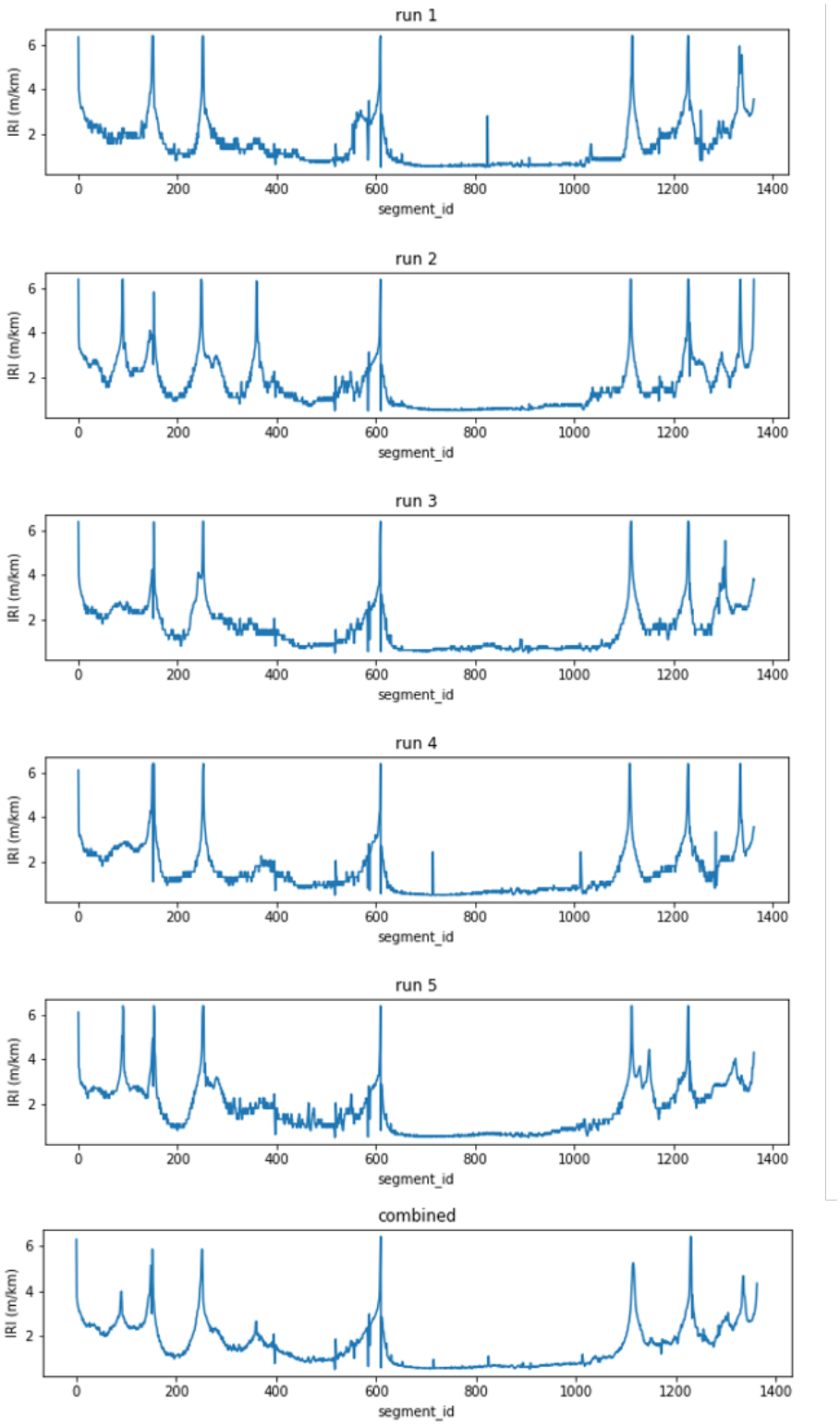



Figure 5.3: IRI estimated from individual runs and after multiple run pavement condition estimation

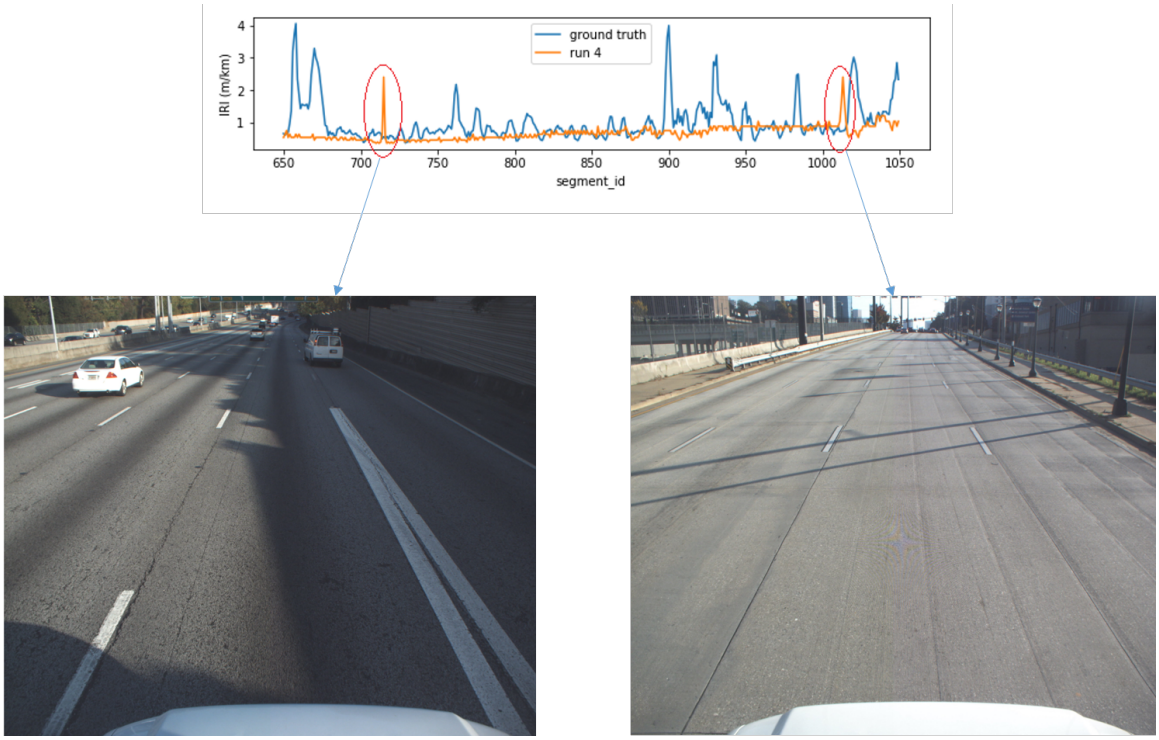



Figure 5.4: Sudden spikes in IRI estimated from a single run

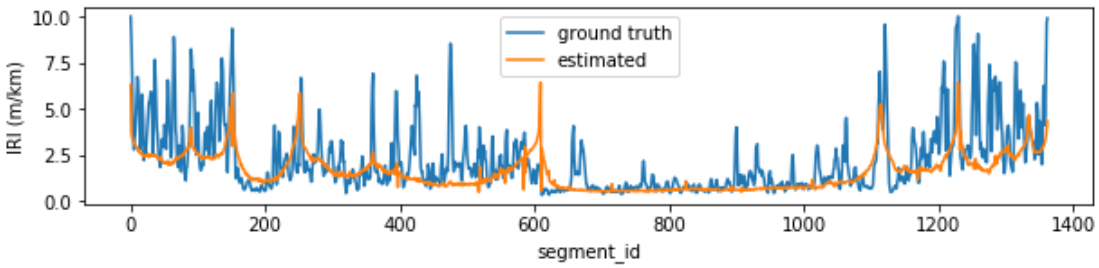



Figure 5.5: Estimated and ground truth IRI on GT Test Route

signed for each pavement condition estimate, as explained in section 5.1, for the simulated crowdsourced runs used in section 5.3.1.

The confidence level results along the GT Test Route are shown in figure 5.6. The locations with the worst confidence levels were qualitatively assessed. The confidence level lowered at lane changes (figure 5.7) as well as roads with bad condition (figure 5.8). Lower confidence level at lane changes are expected as different runs may capture different lanes with naturally different conditions. A lower confidence level at very bad road conditions indicates the difficult cases for the proposed methodology, where it might be useful to reject the pavement condition estimates from smartphone data.

To quantitatively validate the utility of the confidence level, the actual RMSE in determining the IRI using combined smartphone data was calculated using 3D pavement data for the same road segments collected at the same time step. A desirable negative linear correlation was found between the confidence level and RMSE, as shown in figure 5.9. Absolute error was used in the y-axis to more clearly show the overall trend in the presence of outliers. Choosing a cutoff value for the confidence level below which the pavement condition estimate will be rejected presents a tradeoff. If the cutoff is too low, then several bad pavement condition estimates will remain. If the cutoff is too high, too many of the pavement condition estimates will be rejected, diminishing the effort saved for the engineers. Figure 5.10 shows the reduction in RMSE by confidence level cutoff. For example, a confidence level cutoff of 0.75 manages to reduce the RMSE to 1.93 $m/km$, as compared to and RMSE of 2.34 $m/km$ if no segments are removed. Sudden drops in the RMSE appear when large outliers get removed.

The confidence level distribution is shown in figure 5.11. Clearly the confidence level values are concentrated near the higher end of the scale, indicating that high confidence level cutoff values can be used without rejecting too many pavement condition estimates. This is better visualized in figure 5.12, which demonstrates the percentage of pavement condition estimates which will be rejected to achieve a certain reduction in RMSE. It can

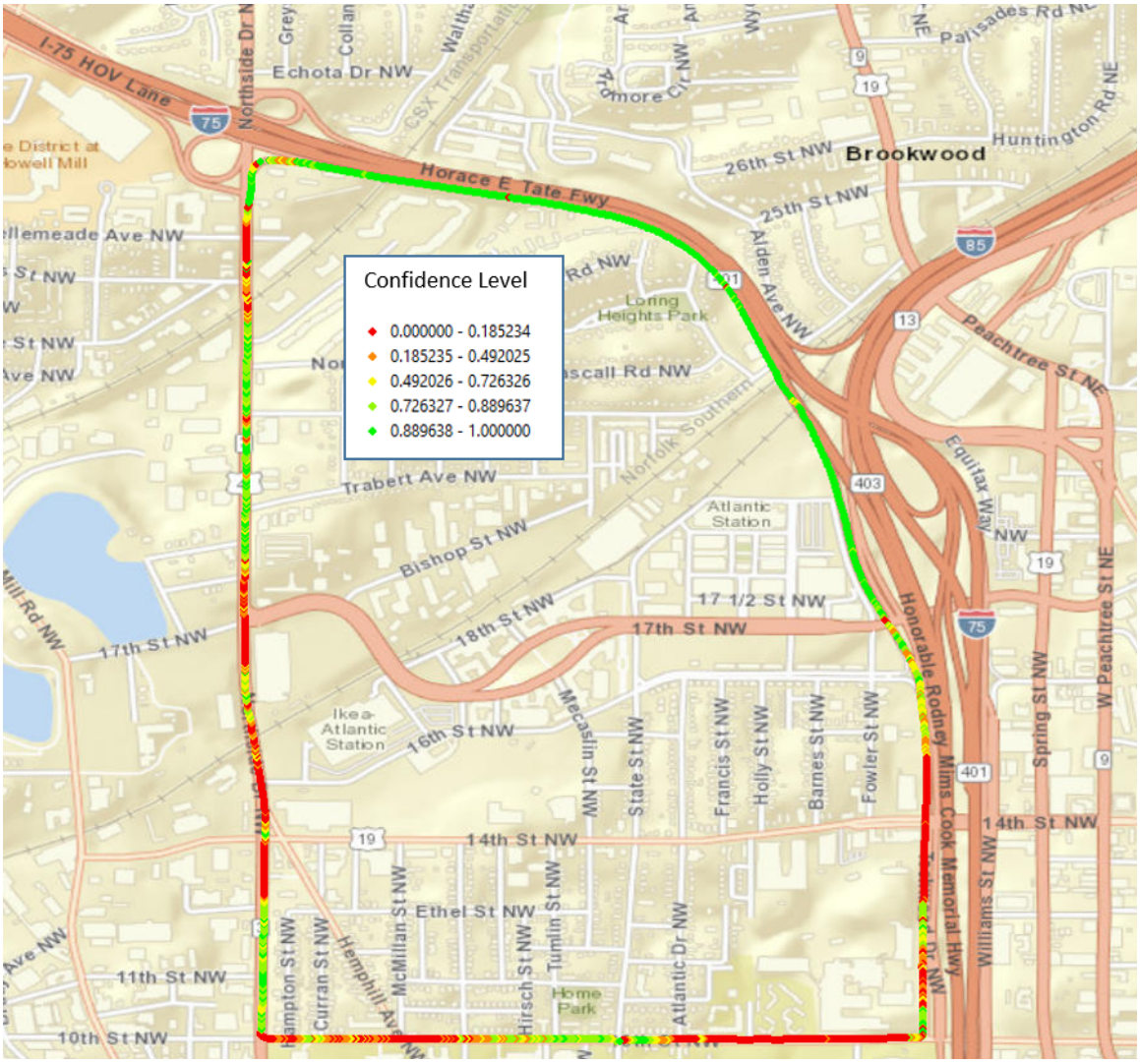



Figure 5.6: Confidence level by variance across runs


Figure 5.7: Low confidence level at lane change

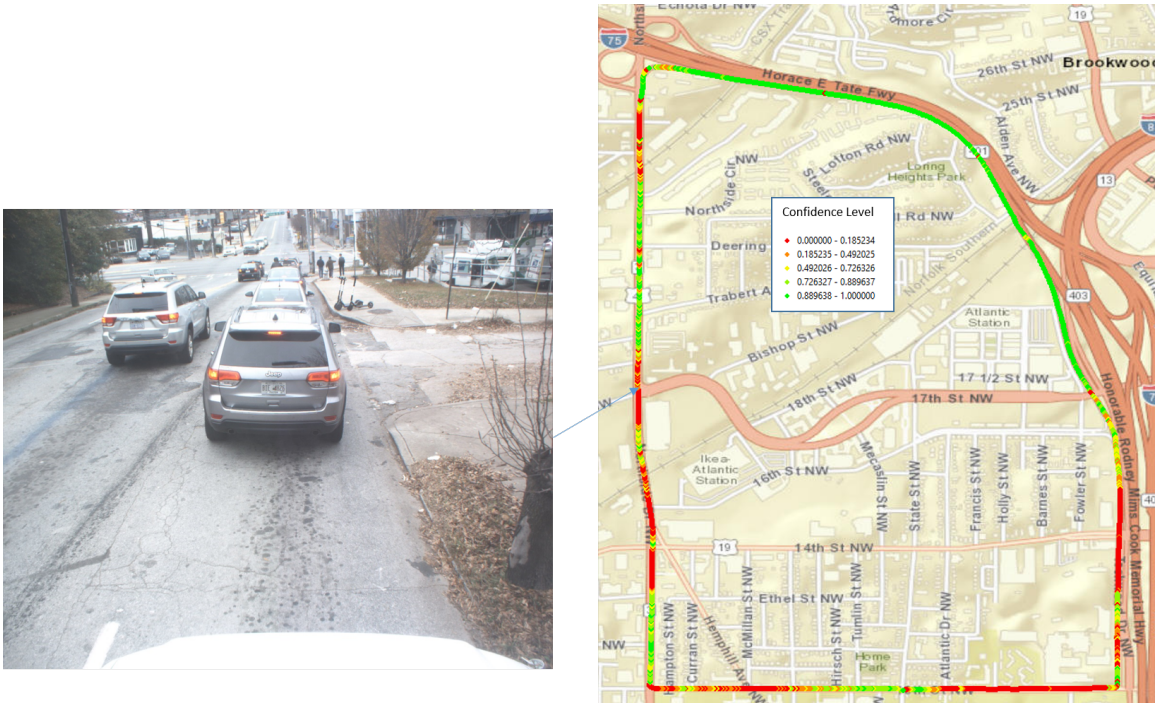



Figure 5.8: Low confidence level at bad pavement condition

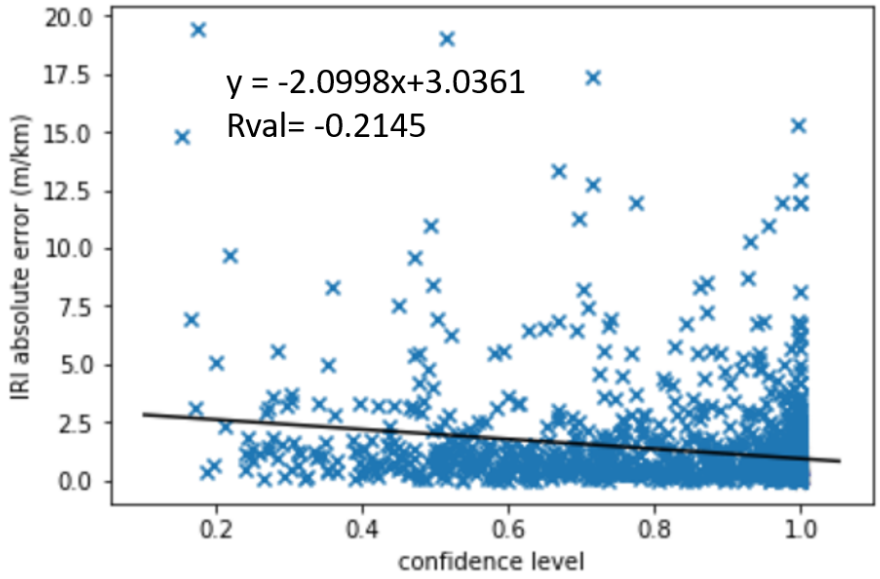



Figure 5.9: Scatter plot of confidence level against RMSE

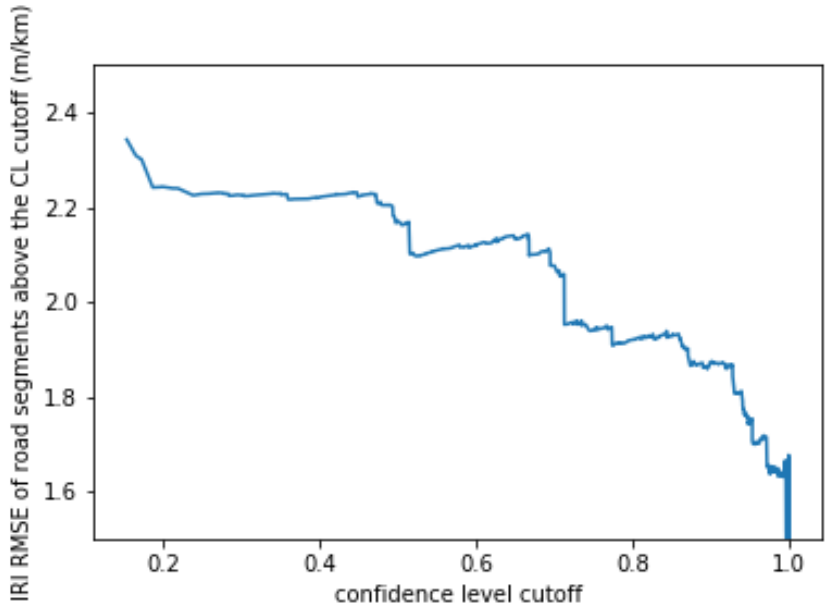



Figure 5.10: RMSE after removing segments below a confidence level cutoff

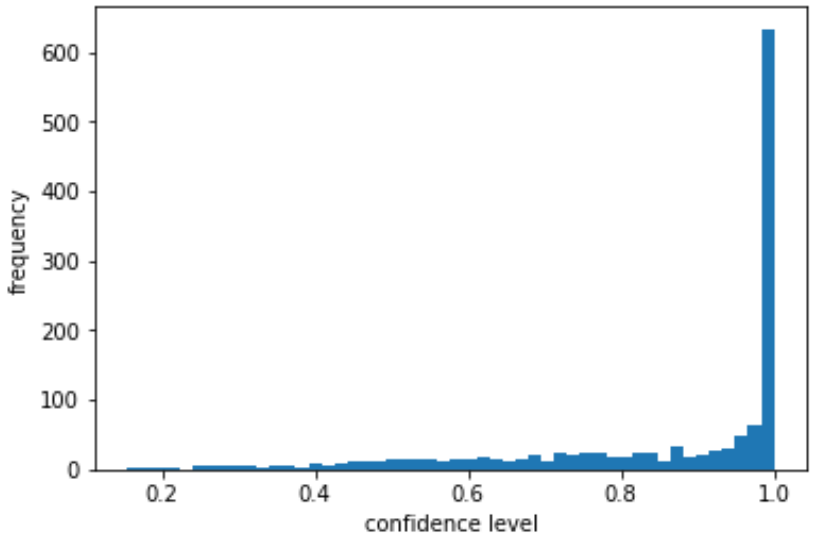



Figure 5.11: Distribution of confidence level values

be observed that removing more than 65% of the road segments is not beneficial, as outliers with high confidence can end increasing the RMSE as more results are removed. Figure 5.13 provides a reference between different confidence level cutoff values and the percentage of road segments that would be removed by that cutoff.

### 5.3.3 Results for proposed methodology for pavement condition prediction

In this thesis, one month was taken as the size of one time step. To validate the proposed methodology for pavement condition prediction presented in section 5.2, 3D pavement data was collected monthly along with smartphone data. Only the 3D pavement data from the initial month was used to provide the initial prior pavement condition information. For subsequent months, only the smartphone data was provided as evidence to update the parameters defining the pavement condition. The remaining 3D pavement data was used to quantitatively assess the performance of the proposed methodology (figure 5.14).

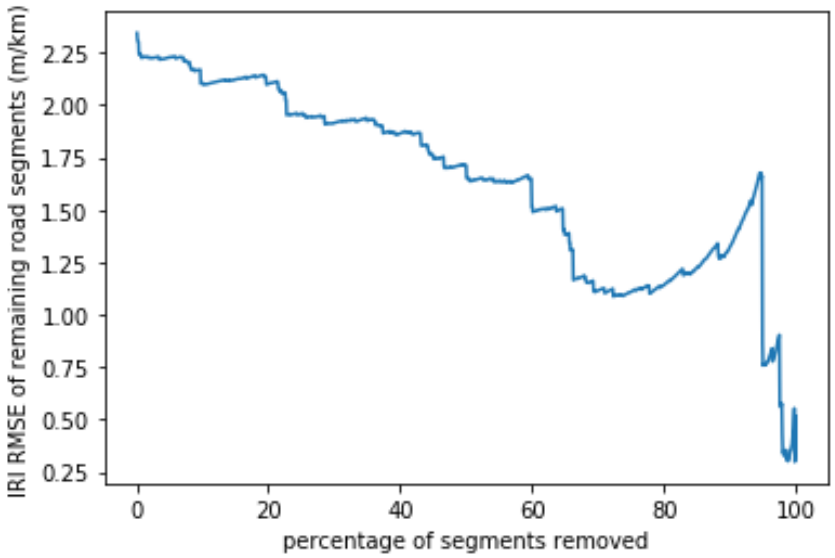



Figure 5.12: RMSE after removing the lowest confidence level road segments

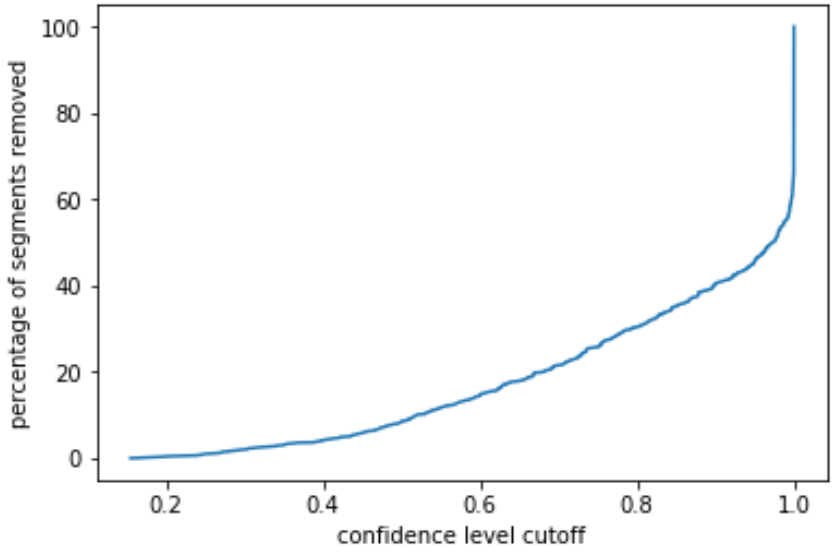



Figure 5.13: Percentage of road segments below confidence level cutoffs

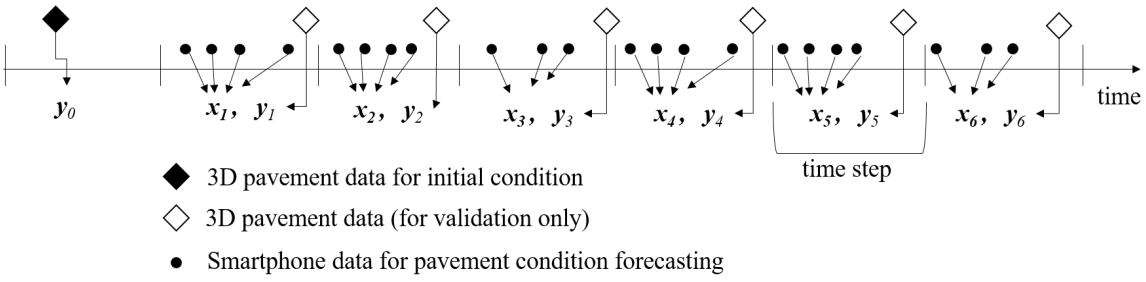



Figure 5.14: Illustration of use of data in pavement condition forecasting

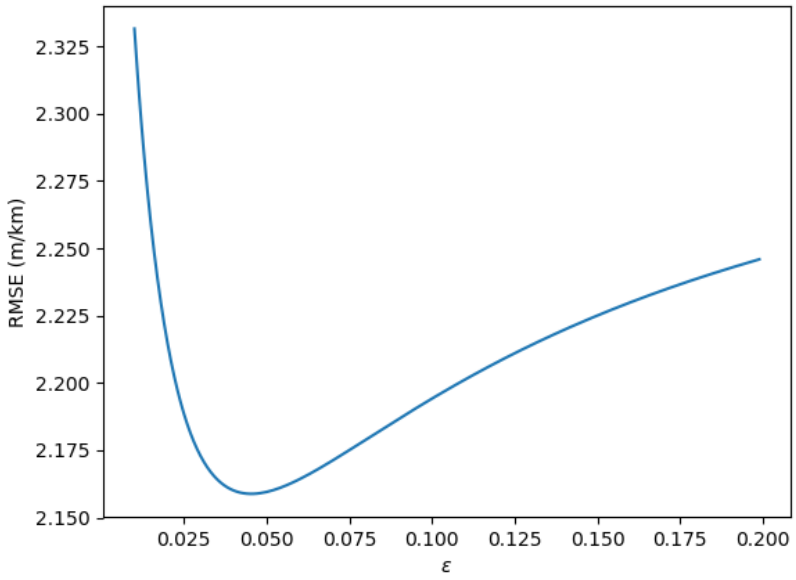



Figure 5.15: Hyperparameter $\epsilon$ vs RMSE

*Optimal value of hyperparameter* $\epsilon$

The hyperparameter $\epsilon$ was introduced in section 5.1 to provide different weights to pavement condition estimates from smartphone and 3D pavement data. If the values of $\epsilon$ is too high, the 3D pavement data, which gives the initial prior pavement condition, will receive too much weight and collected smartphone evidence at the first time step will have little effect on updating the pavement condition. If $\epsilon$ is too low, the prior information provided by the highly accurate 3D pavement data will be discarded too easily in the first time step.

The value of $\epsilon$ will depend on the distress value being estimated. To find the optimal value for the hyperparameter $\epsilon$ for IRI, the RMSE of predicting the IRI value of the first time step was plotted against different values for $\epsilon$, as shown in figure 5.15. An optimal value for $\epsilon$ was determined to be 0.0450.

*Optimal value of initial prior values of parameter* $\alpha$

The hyperparameter $\alpha_0$ defines the initial distribution of the random variable $\sigma^2$ (equation 5.16), which gives the variance of the likelihood (equation 5.14). We have a realization

Table 5.1: IRI RMSE by approach

| | IRI RMSE (m/km) |
|---|---|
| No Prior | 2.3430 |
| Variance Known | 2.1588 |
| Variance Unknown | 2.1588 |

of this random variable for every segment collected during the simulated crowdsourcing drawn from the distribution of random variable $\sigma$. Thus, the obtained set of realizations can be fit to an inverse gamma distribution to approximate the parameter $\alpha_0$. A maximum likelihood estimate of 0.6299 was obtained for the hyperparameter $\alpha_0$.

*Validation of proposed methodology for pavement condition prediction*

With the optimal fixed parameters for the initial prior distribution, the RMSE was determined for both variance known and variance unknown cases and compared with the base case of pavement condition estimation without using prior information. An improvement in the IRI RMSE was observed when prior information is used, indicating that the prior information helped to steer the posterior pavement condition estimate towards the correct value. However, the difference between the variance known and variance unknown models was negligible.

# CHAPTER 6

# CASE STUDIES

## 6.1 Validation of IRI Estimation Using 3D Pavement Data

On February 14, 2019, our research group collected pavement condition data on two test sections (one asphalt and one concrete) using the GTSV and GDOT Profilers. The GDOT profilers are certified for pavement roughness estimation (including IRI) according to the certification procedure in AASHTO R56 (AASHTO, 2014). The GDOT profilers (figure 6.1) comprise of a system of 7 laser sensors affixed to the front bumper of a vehicle. The laser sensors capture a thin longitudinal profile of the pavement surface each, which is used to determine IRI.

The asphalt test section was located inside Middle Georgia Regional Airport near Macon, GA (figure 6.2b) and the concrete test section was located on U.S. Route 41 near Barnesville, GA (figure 6.2a). The test sections consisted of a 0.1 mile (528 feet) stretch for which the IRI was to be determined, with 500 feet buffers on each side. The test section was marked with reflective tape, as shown in figure 6.2b, which be easily visible in both the GDOT profiler's longitudinal profiles as well as the range images from the GTSV's laser


Figure 6.1: GDOT Profiler

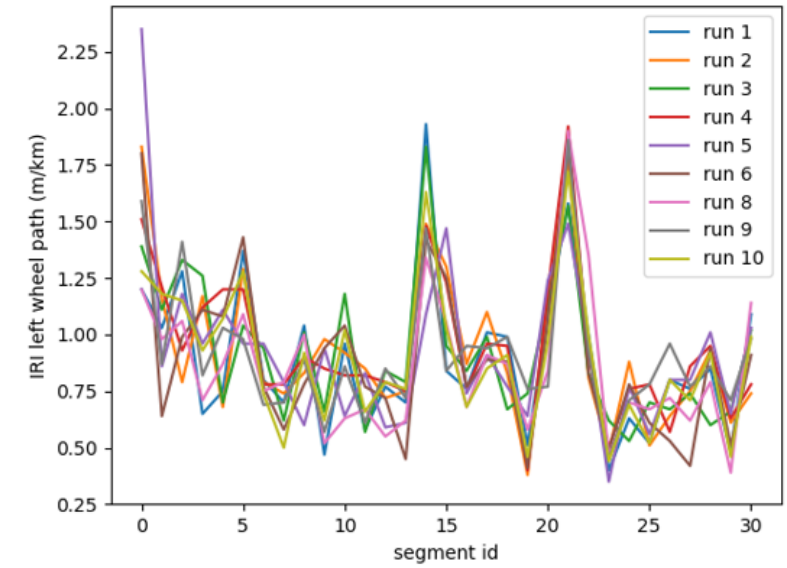
(a) (b)

Figure 6.2: IRI Test sections in (a) U.S. 41 (Concrete) and (b) Middle Georgia Regional Airport (Asphalt)

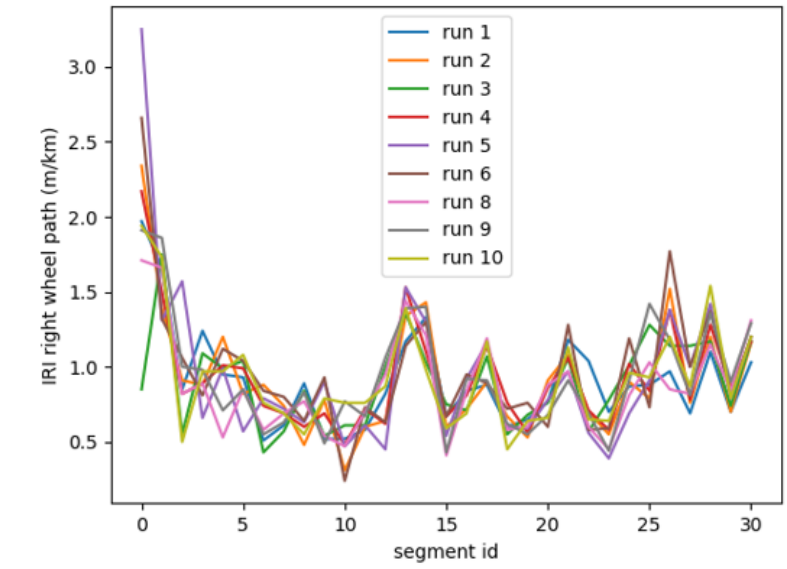
scanner. The buffers were marked with traffic cones (figure 6.2a). Each test section was covered in multiple runs by both the GDOT Profiler and the GTSV.

First, the repeatability of the IRI estimated from data collected using the GTSV was evaluated. Each test section was collected in 10 runs using the GTSV. The runs were manually registered to align them by locating the reflective strip at the start of the test section in the intensity images for each run. The IRI was then estimated using the vendor provided software LCMS RoadInspect. LCMS RoadInspect estimates the IRI for each 5 meter segment. The IRI estimated from each run for the left and right wheelpaths from the asphalt test section is given in figure 6.3. Run 7 was omitted as the vehicle path was askew during that run. Qualitatively, the repeatability of the estimated IRI can be observed especially in the features at segments 13-23 in the left wheel path and segments 13-22 in the right wheel path. Quantitatively, the Pearson correlation (equation 3.2) between each series of IRI values derived from each run was calculated. As shown in figure 6.4, there is

generally a high correlation between any two runs, with an average correlation of 0.7912 for the left wheel path and 0.7876 for the right wheel path.



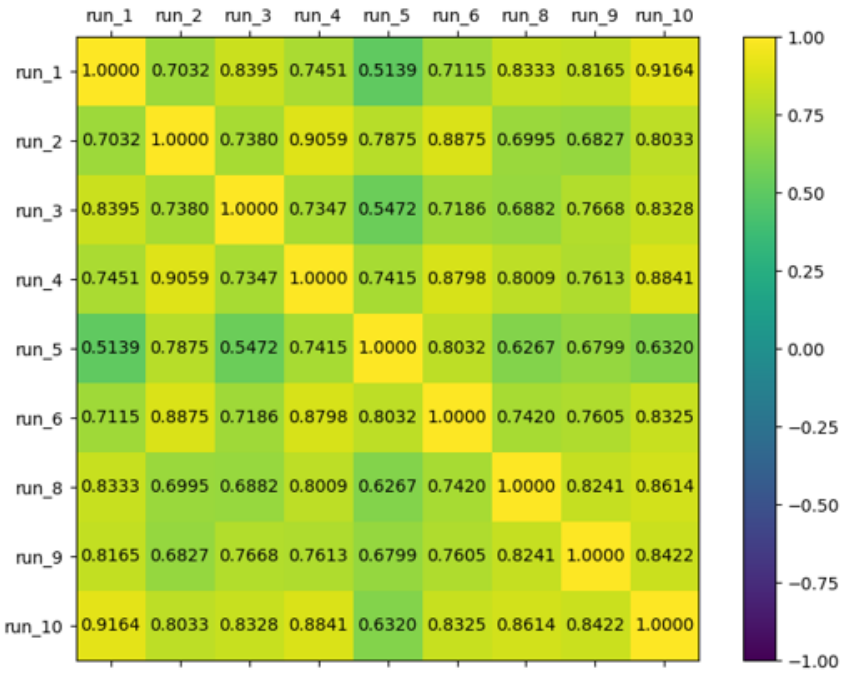

(a)




(b)

Figure 6.3: IRI estimated using GTSV data on (a) left wheel path and (b) right wheel path

Second, the accuracy of IRI estimation using the GTSV was evaluated. GDOT provided the longitudinal profiles collected using the GDOT profilers, which were used to determine




(a)


(b)

Figure 6.4: Correlation between IRI estimated from different GTSV runs on (a) left wheel path and (b) right wheel path

the IRI for each wheelpath individually using the *ProVal* software. Similarly, the IRI was estimated using runs of data collected using the GTSV. The calculated IRI from each device is tabulated in table 6.1. The difference between the average IRI estimates is -0.58 $in/mi$ for the left wheel path and 2.57 $in/mi$ for the right wheel path. To put these values in context, the ride quality classifications for different IRI values as defined by the Federal Highway Administration (FHWA) and New York State Department of Transportation (NYSDOT) are given in figure 6.5. We can observe that the slab size for the classification levels is much larger than the differences in IRI values estimated by the profiler and GTSV. Thus, the GTSV provides an accurate estimate of the pavement IRI for the purposes of pavement condition assessment and maintenance.

Table 6.1: IRI determined by GDOT Profiler and GTSV

| run | GDOT IRI_left $(in/mi)$ | GDOT IRI_right $(in/mi)$ | GTSV IRI_left $(in/mi)$ | GTSV IRI_right $(in/mi)$ |
|---|---|---|---|---|
| 1 | 62.16 | 61.13 | 62.07365 | 62.57285 |
| 2 | 63.06 | 62.01 | 61.38005 | 61.59785 |
| 3 | 63.13 | 60.86 | 63.62885 | 64.16645 |
| 4 | 63.26 | 60.63 | 64.41605 | 65.72166 |
| 5 | 62.87 | 61.37 | 61.91465 | 64.82526 |
| 6 | 63.35 | 61.86 | 60.92165 | 64.37765 |
| Average | 62.97 | 61.31 | 62.38915 | 63.87695 |

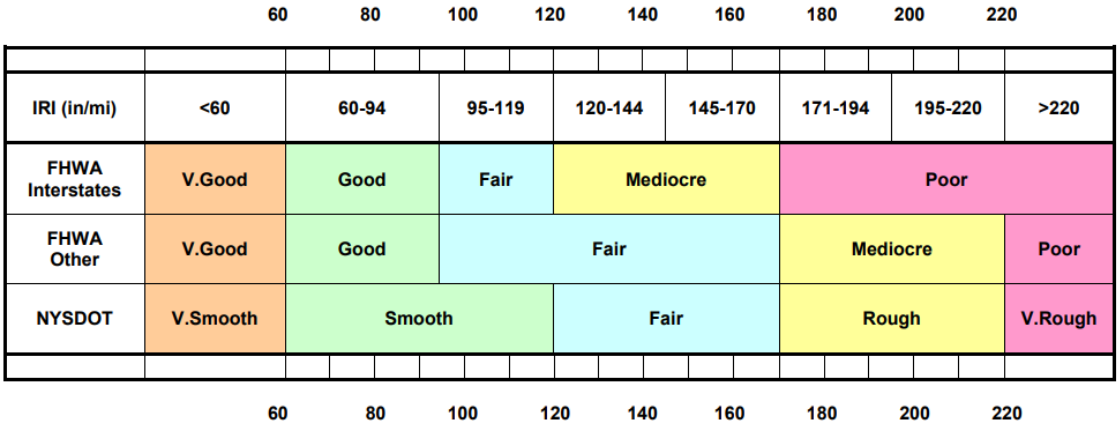



Figure 6.5: Ride Quality Classification by FHWA and NYSDOT

Table 6.2: Distribution of data collection route by road condition

| Condition | Road | Length (mi) |
|---|---|---|
| Bad | 10th St NW + W Wesley Rd NW | 5.6 |
| OK | Northside Dr + Exit 250 + Techwood Dr | 2.6 |
| Good | I-75 NB + SB + I-285 CW | 9.7 |
| Total | | 17.9 |

Table 6.3: Distribution of data collection route by road classification

| Classification | Length (mi) |
|---|---|
| Local | 4.7 |
| Arterial | 1.5 |
| Freeway | 9.7 |

## 6.2 Analysis of Data Collection Route

To ensure that multiple pavement types and conditions are considered, the following route was used for data collection (Fig. 6.6). The proposed route covers local streets (W Wesly Rd NW), arterials (Northside Dr NW) and interstates (I-75 NB/SB). Multiple pavement conditions are represented by this route, as shown in figure 6.7. Additionally, to cover different pavement types (dense graded asphalt, OGFC asphalt and concrete), a concrete section of I-285 was also collected. The distribution of road conditions and classifications collected are summarized in tables 6.2 and 6.3.

### 6.2.1 Data Collection

Road infrastructure condition data was collected monthly using the GTSV to provide an accurate estimation of the road condition at different timestamps. This data served as the high confidence or ground truth road condition for training road condition estimation models using smartphones. To simulate crowdsourcing of data, multiple runs using different vehicles and smartphone devices were collected, clustered to multiple timestamps. To collect a large number of runs on the same day, a shortened version of the data collection route, which still covered different pavement condition and road classifications was used. This

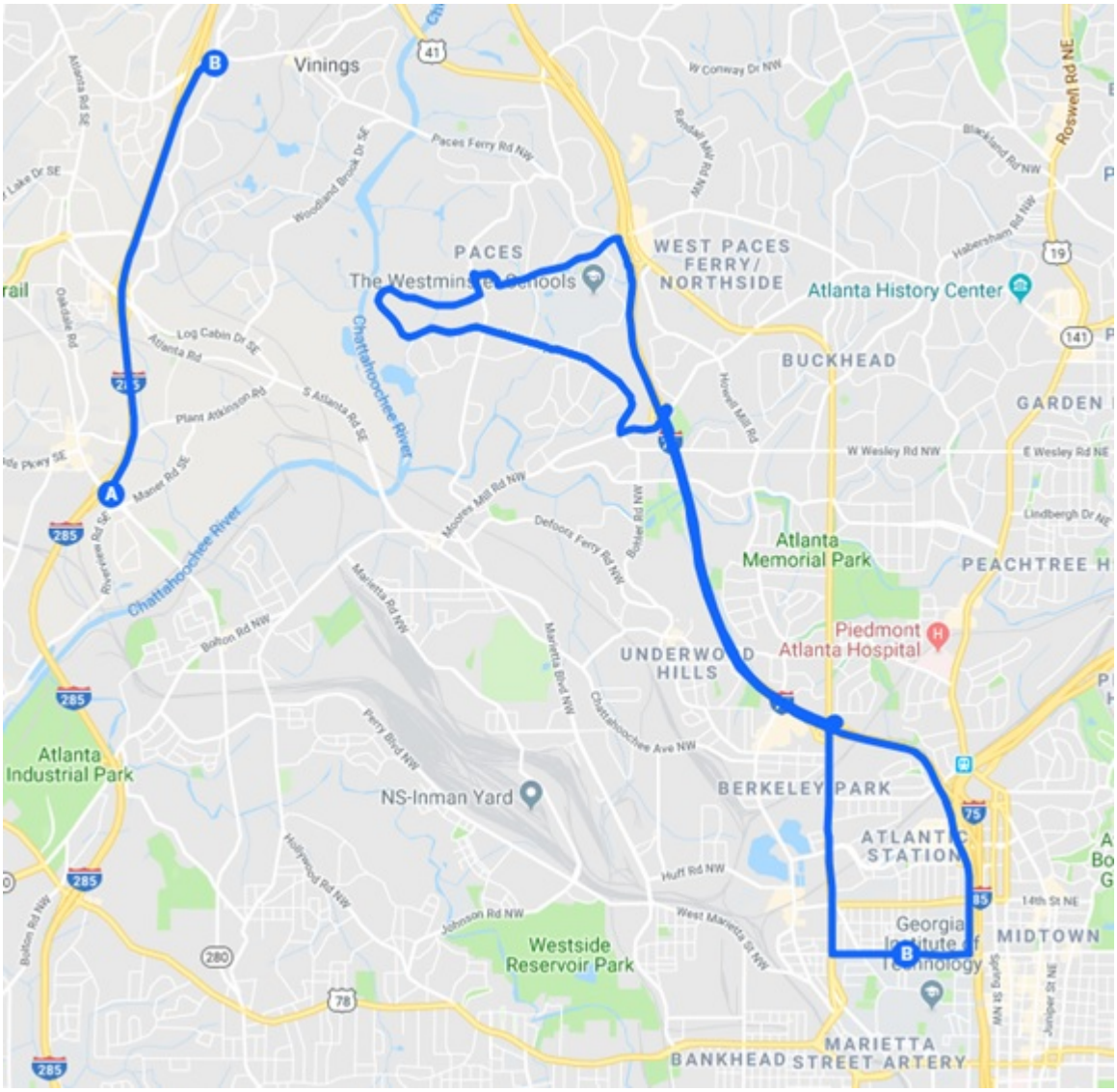



Figure 6.6: Data Collection Route


Figure 6.7: Video data from data collection route showing: a) and b) patches on 10th St c) Cracking on W Wesley Rd d) Alligator cracking on W Wesley Rd

Table 6.4: Distribution of GT test route by road condition

| Condition | Road | Length (mi) |
|---|---|---|
| Bad | 10th St NW | 0.70 |
| OK | Northside Dr + Exit 250 + Techwood Dr | 2.03 |
| Good | Hemphill Ave + I-75 SB | 1.39 |
| Total | | 4.12 |

route, referred to as the GT Test Route, is shown in figure 6.8 and its breakdown by pavement condition and road classification is given in tables 6.4 and 6.5 respectively. Each of these timestamps corresponded with one of the periodic sensing van runs, which provided the ground truth road condition for that time. The crowdsourced data was used to estimate the road condition using multiple run data (section 5.1). Table 6.6 lists all data collection runs. As explained in section 3.2, each run in the main data collection route covered 4881 5-meter segments while the GT Test Route consisted of 1366 5-meter segments.

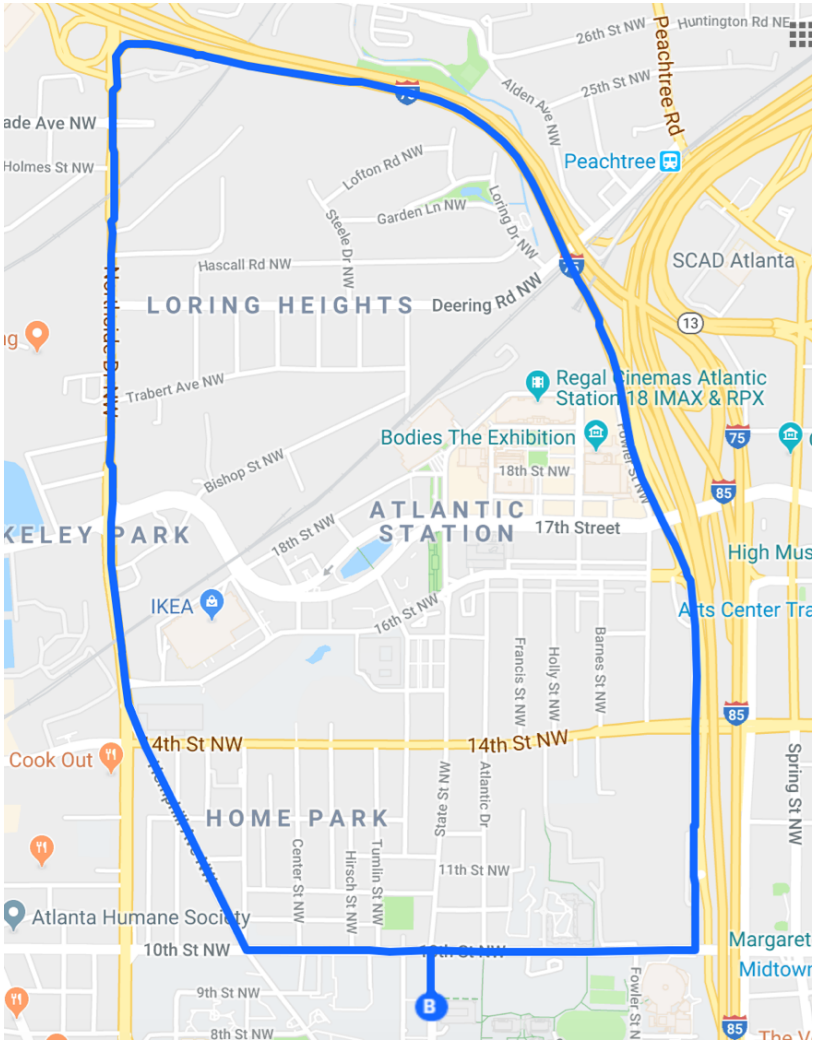



Figure 6.8: GT Test Route

Table 6.5: Distribution of GT test route by road classification

| Classification | Length (mi) |
|---|---|
| Local | 1.13 |
| Arterial | 2.03 |
| Freeway | 0.96 |

Table 6.6: Data Collection Runs

| S. No. | Run ID | 3D Pavement Data | Smartphone Data |
|---|---|---|---|
| 1 | 20171024 Run 1 | Y | Y |
| 2 | 20171024 Run 2 | Y | Y |
| 3 | 20171024 Run 3 | Y | Y |
| 4 | 20180308 Run 1 | Y | N |
| 5 | 20180308 Run 2 | Y | Y |
| 6 | 20180321 Run 1 | N | Y |
| 7 | 20180321 Run 2 | N | Y |
| 8 | 20180321 Run 3 | N | Y |
| 9 | 20180321 Run 4 | N | Y |
| 10 | 20180321 Run 5 | N | Y |
| 11 | 20181021 Run 1 | Y | Y |
| 12 | 20181119 Run 1 | Y | Y |
| 13 | 20181213 Run 1 | Y | Y |
| 14 | 20181216 Run 1 | N | Y |
| 15 | 20181216 Run 2 | N | Y |
| 16 | 20181216 Run 3 | N | Y |
| 17 | 20181216 Run 4 | N | Y |
| 18 | 20181216 Run 5 | N | Y |
| 19 | 20190105 Run 1 | Y | Y |
| 20 | 20190320 Run 1 | Y | Y |

### 6.2.2 Analysis of Pavement Distresses in Data Collection Route

The pavement distresses were quantified as described in section 4.1.3. Figure 6.9 shows the distribution of distress values on the route for one run (20181021 Run 1). The total number of segments is 4881. Notably, we observe that the IRI and raveling index have a small number of large values in their range, and a very small percentage of road segments have potholes (0.98%). The outliers in the IRI, rutting and raveling values are more clearly visible in the box and whisker plot in figure 6.10, which are all large values. The frames corresponding to extremely large values were investigated manually, and found to occur at heavily deteriorated road segments as shown in figure 6.11. Removing the highest 1% of values in IRI, rutting and raveling helps to remove some of the outliers, giving a better description of the their distribution (figure 6.12).

The scatter plots (figure 6.13) and correlation matrix (figure 6.14) between the different distresses show minimal correlation between them. The best correlation is in fact a negative correlation of $-0.377$ between raveling index and the maximum crack level. The highest positive correlation is between IRI and maximum crack level (0.327).

## 6.3 Analysis of Proposed Methodology for Single Run Pavement Condition Estimation

The methodology for single-run pavement condition estimation using smartphone data presented in chapter 4 was applied on a run of the data collection route which was not used for training for estimating IRI. The 3D pavement data for that run was used to establish the ground truth pavement condition against which the returned IRI values were compared. The distribution of IRI error by pavement condition and road classification is given in table 6.7 and table 6.8 respectively. It can be observed that the model performs significantly better on good road conditions. By looking at the distribution of the RMSE values for the segments, it can be observed that most of the errors are skewed towards the lower end in

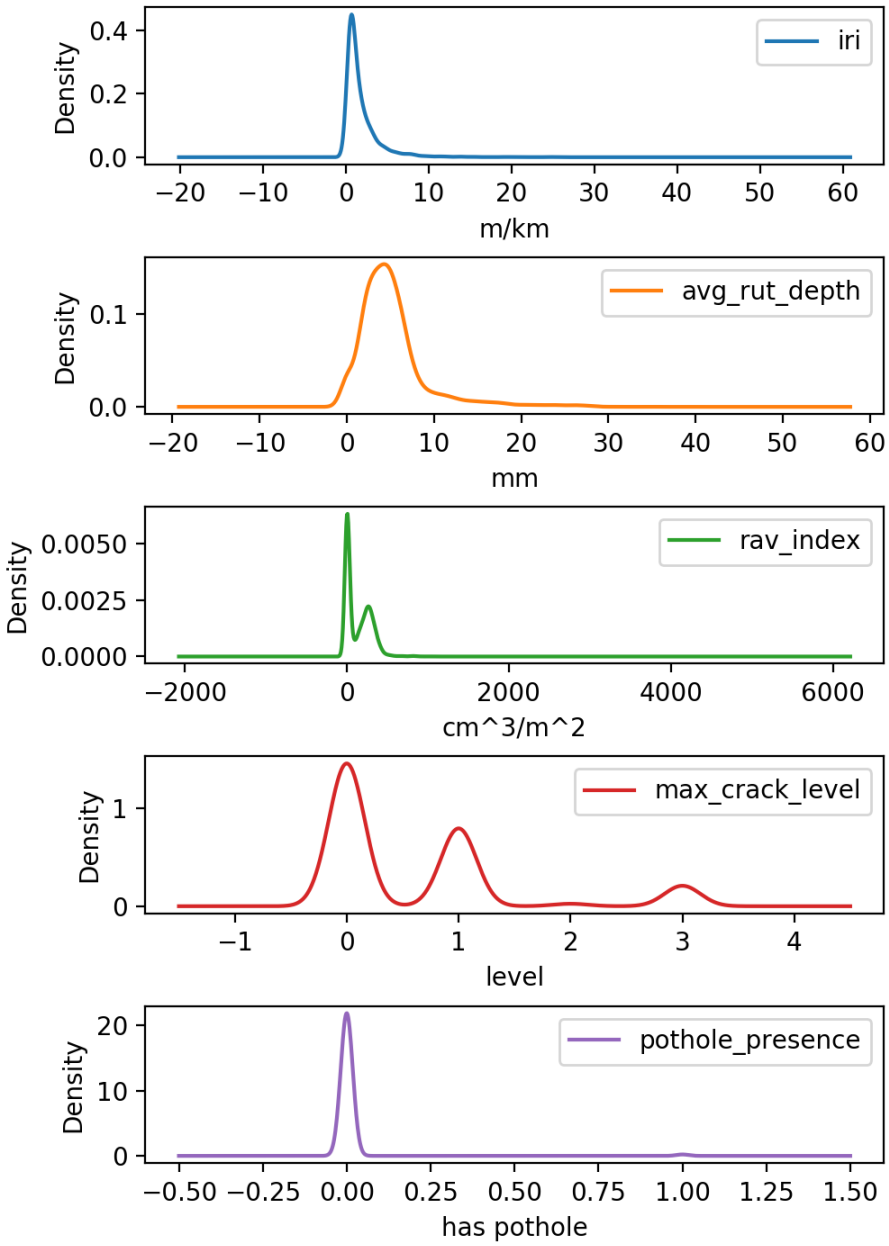



Figure 6.9: Distribution of Distress values

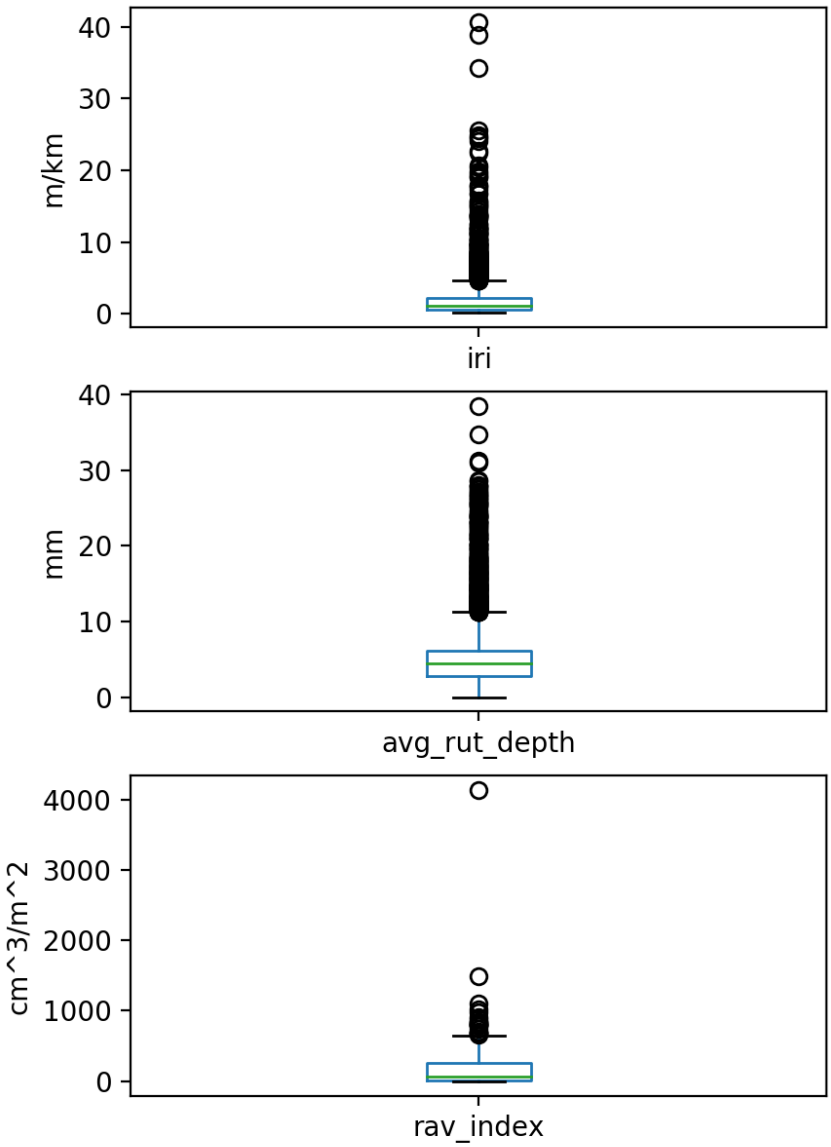



Figure 6.10: Boxplot of Distress values

(a) (b)

Figure 6.11: Camera image of road segments with (a) high IRI value (b) high raveling index value

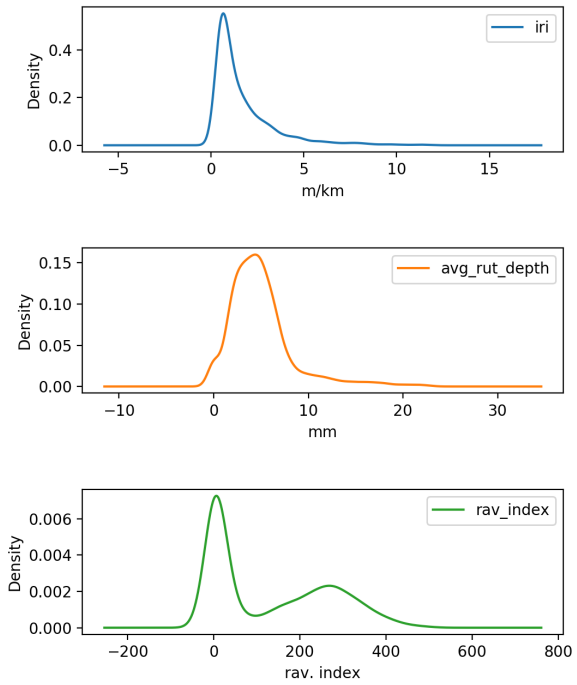



Figure 6.12: Distribution of distress values after removing highest 1% values

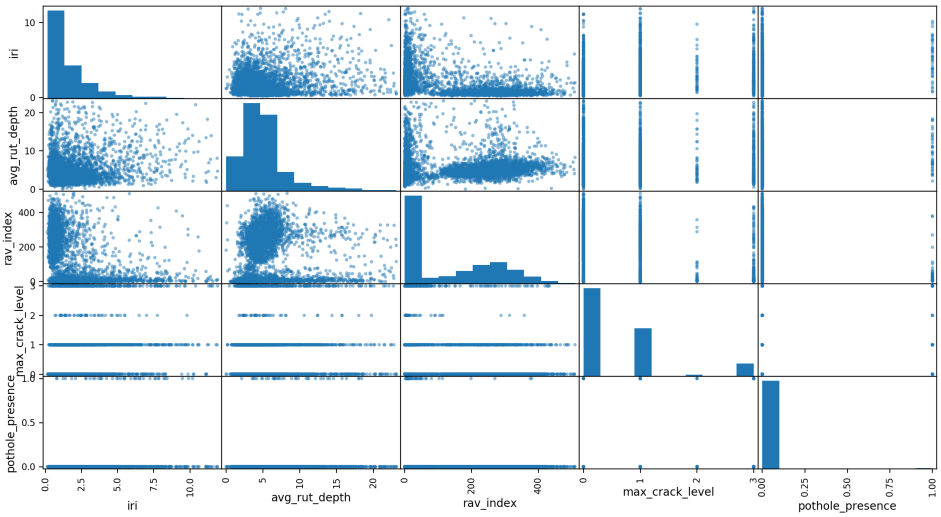



Figure 6.13: Scatter plot matrix of distress values

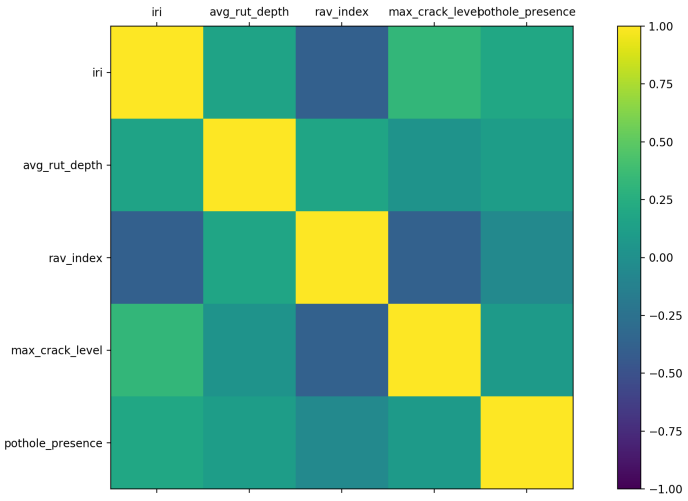
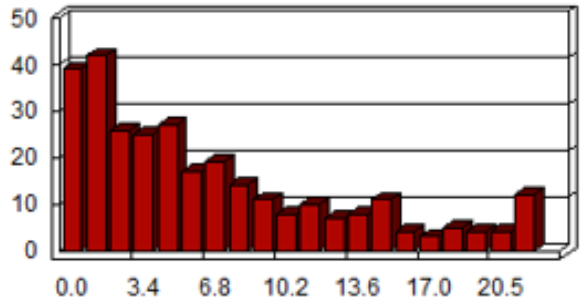
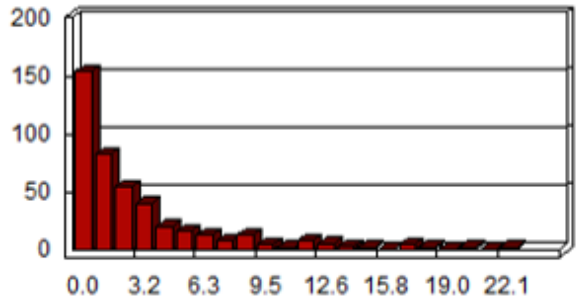
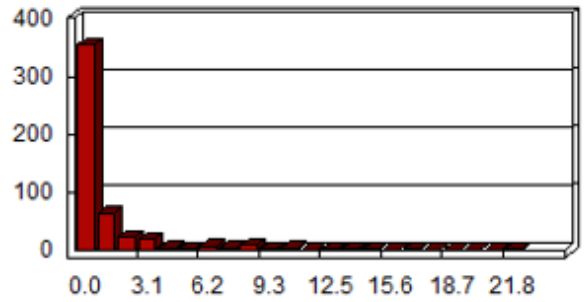



Figure 6.14: Correlation matrix of distress values

Table 6.7: IRI RMSE by pavement condition using single-run pavement condition estimation

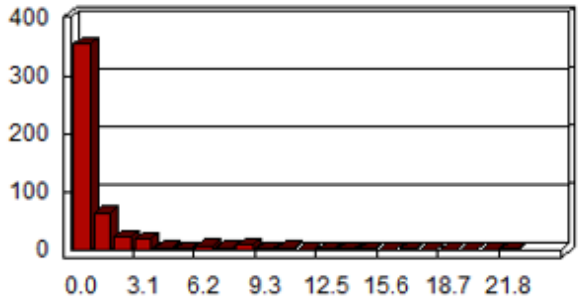
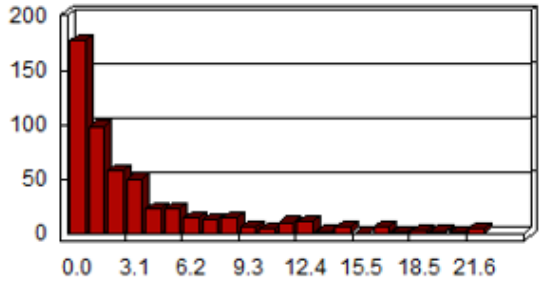
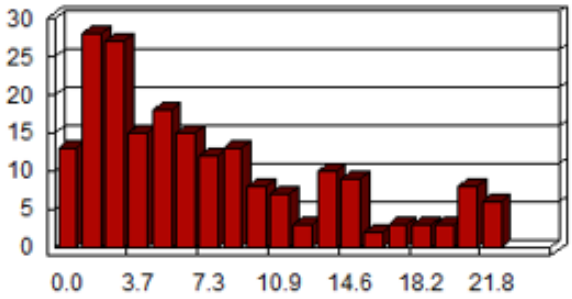
| Pavement Condition | IRI RMSE (m/km) | Distribution of RMSE (IRI RMSE (m/km) vs frequency) |
|---|---|---|
| Bad | 2.67 |   |
| OK | 1.81 |   |
| Good | 1.19 |   |

all cases, with a small number of high error values, which could lead to a high overall RMSE. Thus, it is important to detect and reject bad pavement condition estimates, which was explored in chapter 5.

## 6.4 Analysis of Proposed Methodology for Multiple Run Pavement Condition Estimation

The methodology for multiple-run pavement condition estimation using smartphone data presented in chapter 5 was applied on runs 14-18 of the data collection route which was

Table 6.8: IRI RMSE by road classification using single-run pavement condition estimation

| Road Classification | IRI RMSE (m/km) | Distribution of RMSE (IRI RMSE (m/km) vs frequency) |
|---|---|---|
| Freeway | 1.19 |   |
| Arterial | 1.91 |   |
| Local | 2.81 |   |

not used for training for estimating IRI. A combined IRI estimate for the road segments was obtained. The 3D pavement data from run 13 was used to establish the ground truth pavement condition against which the returned IRI values were compared. The distribution of IRI error by pavement condition and road classification is given in table 6.9 and table 6.10 respectively. Again, it can be observed that the model performs significantly better on good road conditions. The distribution in all cases are be highly skewed left, indicating that the RMSE may underestimate the performance of the algorithm: a vast majority of road segments actually have a much lower error than the RMSE.

## 6.5 Analysis of Proposed Methodology for Pavement Condition Prediction

The methodology for pavement condition prediction using smartphone data presented in chapter 5 was used to predict the IRI on the data collection route over the monthly collected data. The strategy explained in section 5.3.3 was followed: both 3D pavement and smartphone data was collected monthly. However, only the first month's 3D pavement data and subsequent months' smartphone data was used for pavement condition prediction. The remaining 3D pavement data was only used to assess the performance.

Using the model with known likelihood variance, the distribution of IRI error by pavement condition and road classification is given in table 6.11 and table 6.12 respectively. The same for the model with unknown likelihood variance is given in table 6.13 and table 6.14. Both models give very similar performance. It is once again observed that the models perform significantly better on good road conditions with error distributions highly skewed to the left.

## 6.6 Effect of Smoothing

Calculated IRI values are often smoothed over a distance before reporting. Smoothing also reduces the adverse impact of outliers on the error. However, oversmoothing can reduce the granularity of the results. Figure 6.15 plots the root mean squared error of IRI calculated

Table 6.9: IRI RMSE by pavement condition using multiple-run pavement condition estimation

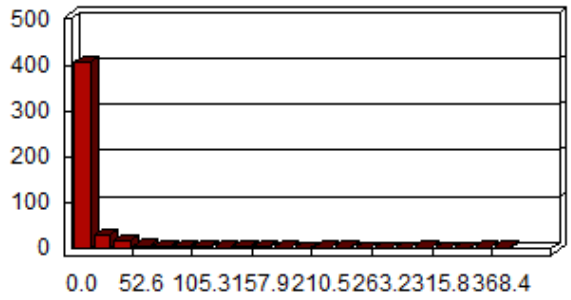
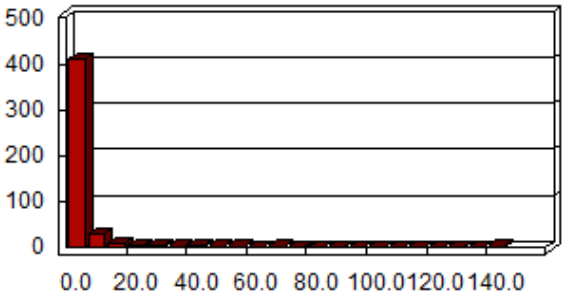
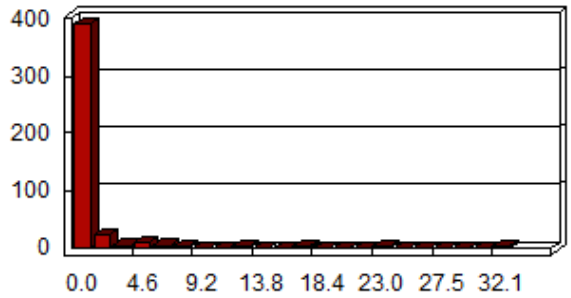
| Pavement Condition | IRI RMSE (m/km) | Distribution of RMSE (IRI RMSE (m/km) vs frequency) |
|---|---|---|
| Bad | 3.5213 |  |
| OK | 1.7110 |  |
| Good | 0.8366 |  |

Table 6.10: IRI RMSE by road classification using multiple-run pavement condition estimation

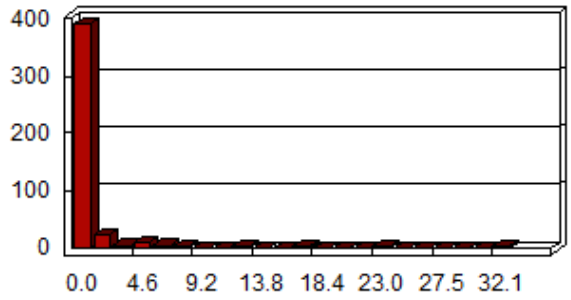
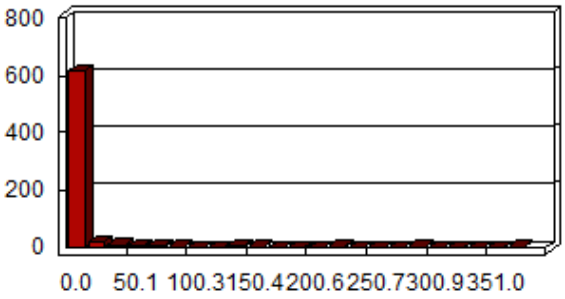
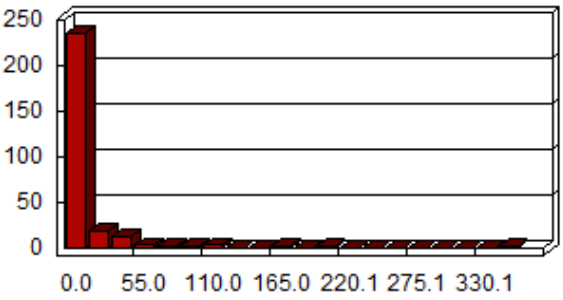
| Road Classification | IRI RMSE (m/km) | Distribution of RMSE (IRI RMSE (m/km) vs frequency) |
|---|---|---|
| Freeway | 0.8366 |  |
| Arterial | 2.3987 |  |
| Local | 3.5185 |  |

Table 6.11: IRI RMSE by pavement condition using pavement condition prediction (variance known model)

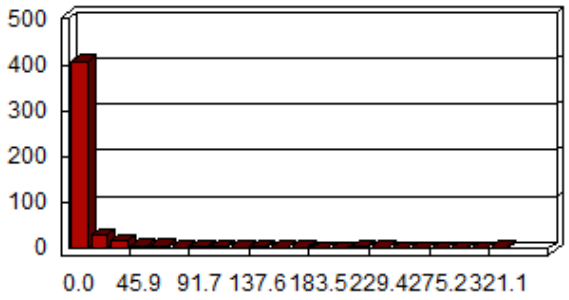
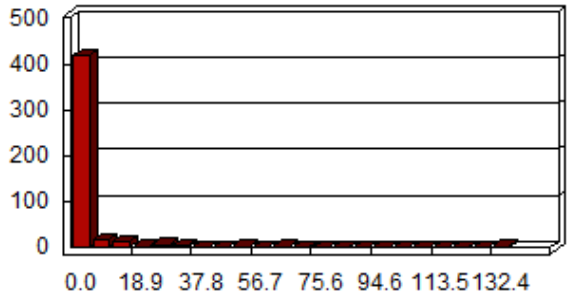
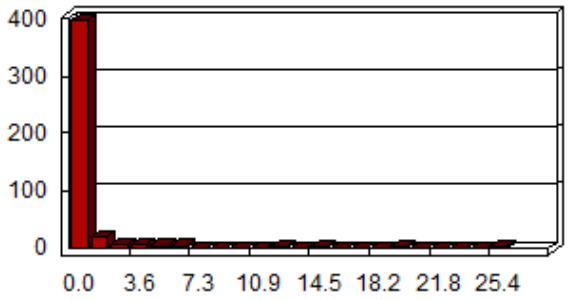
| Pavement Condition | IRI RMSE (m/km) | Distribution of RMSE (IRI RMSE (m/km) vs frequency) |
|---|---|---|
| Bad | 3.2373 |  |
| OK | 1.6099 |  |
| Good | 0.7203 |  |

Table 6.12: IRI RMSE by road classification using pavement condition prediction (variance known model)

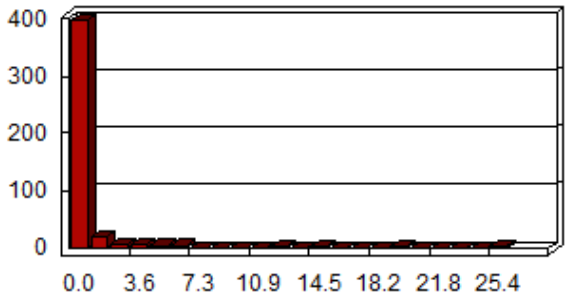
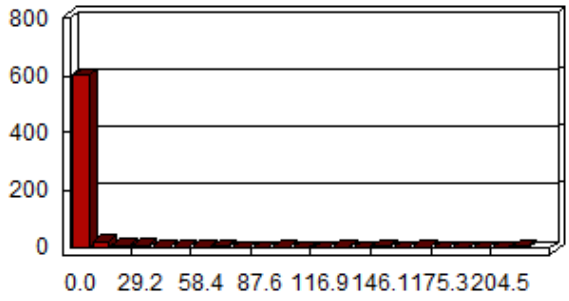
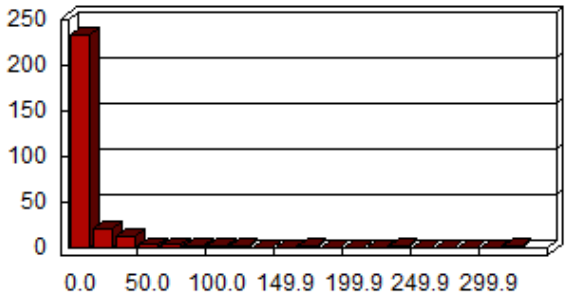
| Road Classification | IRI RMSE (m/km) | Distribution of RMSE (IRI RMSE (m/km) vs frequency) |
|---|---|---|
| Freeway | 0.7203 |  |
| Arterial | 2.1117 |  |
| Local | 3.4098 |  |

Table 6.13: IRI RMSE by pavement condition using pavement condition prediction (variance unknown model)

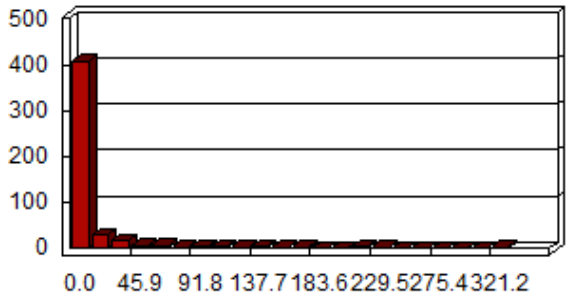
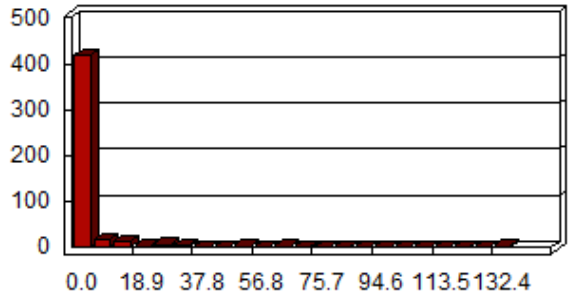
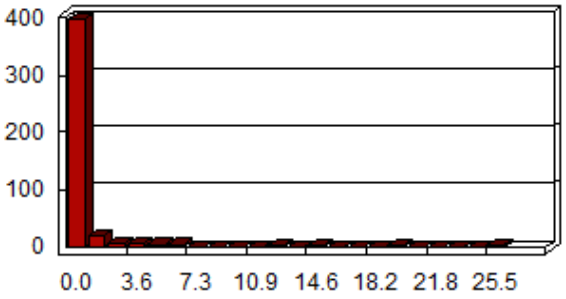
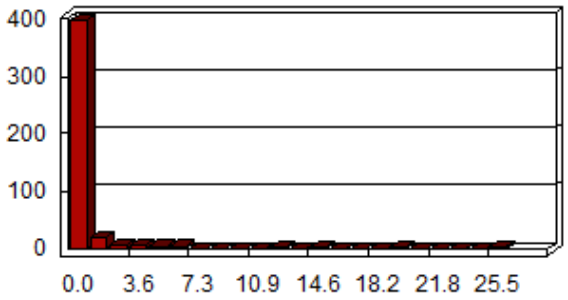
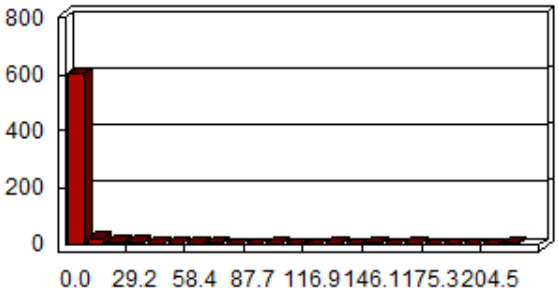
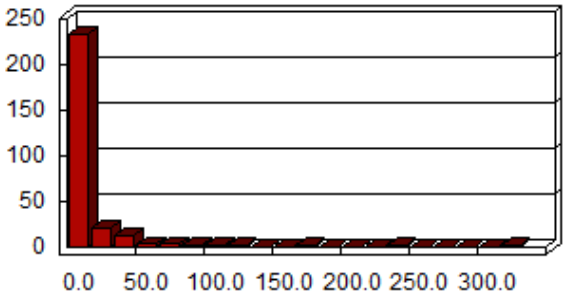
| Pavement Condition | IRI RMSE (m/km) | Distribution of RMSE (IRI RMSE (m/km) vs frequency) |
|---|---|---|
| Bad | 3.2373 |   |
| OK | 1.6097 |   |
| Good | 0.7206 |   |

Table 6.14: IRI RMSE by road classification using pavement condition prediction (variance unknown model)

| Road Classification | IRI RMSE (m/km) | Distribution of RMSE (IRI RMSE (m/km) vs frequency) |
|---|---|---|
| Freeway | 0.7206 |  |
| Arterial | 2.1122 |  |
| Local | 3.4089 |  |

at different levels of smoothing using a moving average filter of varying window sizes. For a series $\{x_1, x_2....x_n\}$, a moving average filter smoothens the series by calculating the average value over a sliding fixed sized window over the series:

$$x'_i = \frac{1}{2b+1} \sum_{k=i-b}^{i+b} x_i \forall i \in \{b+1, ..., n-b\} \tag{6.1}$$

Where $2b+1$ is the window size. It is observed that the RMSE rapidly falls even for small levels of smoothing. The variance known and variance unknown models for pavement condition prediction have very similar RMSE values. Since we have observed that the distribution of errors is highly left skewed, the median absolute error (MAE) can provide a more intuitive idea of the error observed in most segments. This is shown in figure 6.16. The overall trend is a reduction in error for both pavement condition estimation and prediction (both models), with a drastic reduction in MAE for pavement condition prediction.

The actual window size used depends on protocol. IRI is often reported for 0.1 mile (160 meter) segments (FHWA, 2016). Thus, if IRI results are smoothed over 160 meter segments before reporting, the expected RMSE per segment will be 1.13 m/km for multiple-run pavement condition estimation and 0.79 m/km for pavement condition prediction for both models. The median absolute error will be 0.61 m/km for multiple-run pavement condition estimation and 0.36 m/km for pavement condition prediction for both models.

## 6.7 Discussion on Performance

In section 4.2, the performance of the proposed method for single-run pavement condition estimation was quantified using the root mean squared error. However, it was shown that the distribution of errors was left-skewed, so the mean of the squared error will often overestimate the actual error observed by most segments. Furthermore, it was observed that the

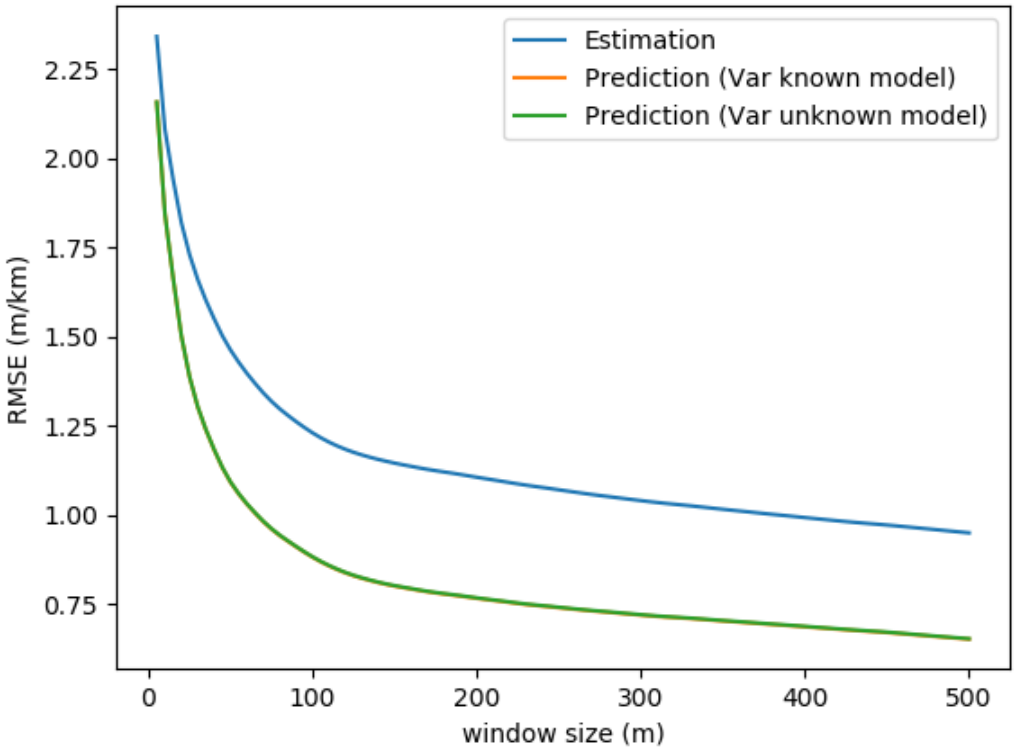



Figure 6.15: Effect of smoothing on IRI RMSE

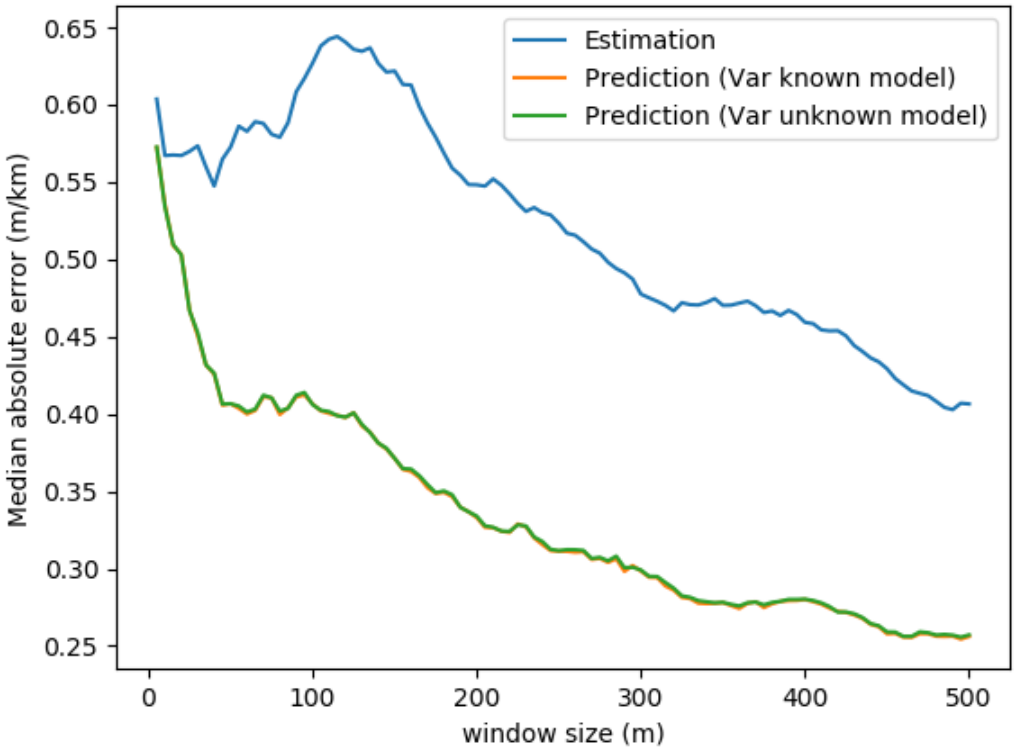



Figure 6.16: Effect of smoothing on IRI median absolute error

high RMSE is heavily reduced after smoothing, which is required in practice as the results are aggregated to a segment size used by different pavement condition assessment protocols. In this section, these factors have been considered to assess the practical performance of the proposed methodologies for pavement condition estimation and prediction.

As discussed in section 6.6, IRI is often reported for 0.1 mile (160 meter) segments, where the median absolute error is 38.65 in/mile for multiple-run IRI estimation and 22.81 in/mi for IRI prediction (both models). On comparing these values to the slab sizes for different pavement condition classifications presented in figure 6.5, it can be observed that the median absolute error is much smaller than the slab sizes, indicating that the proposed methodologies can be successfully applied in the field for improving IRI estimation using smartphones and improving IRI forecasting from prior 3D pavement data.

In the case of rutting, GDOT's PACES protocol (GDOT, 2007) assigns a single rutting deduct value for a 100 ft (30 m) section, which represents the condition for a 1-mile segment. The PACES protocol assigns deduct values according to the average rut depth as per table 6.15. Thus, a different deduct value is assigned for each 1/8 inch = 3.175 mm slab. After smoothing the multiple-run rutting estimates with a window size of 100 ft, the RMSE of rutting values was 4.28 mm while the MAE was 4.19 mm. 36% of the 100 ft segments had rut depths calculated using multiple-run rutting estimation that had an absolute error less than 3.175 mm. For rutting prediction using the known variance model after smoothing, the RMSE dropped to 1.69 mm and MAE dropped to 1.06 mm. 88% of the 100 ft segments now had calculated rut depths that had an absolute error of less than 3.175 mm. For rutting prediction using the unknown variance model after smoothing, the RMSE was 3.04 mm and MAE was 2.89 mm, with 55% of 100 ft segments having an absolute error less than 3.175 mm. Thus, the known variance model is recommended for rutting prediction. Although the error for rutting estimation does not appear satisfactory for practical use, it is still indicative that some information about rutting condition can be captured from vehicle vibration captured using smartphones, which has not been explored before.

Table 6.15: Rutting deduct table (GDOT, 2007)

| Average rut depth (inches) | Average rut depth (mm) | Deduct Value |
|---:|---:|---:|
| 0/8 | 0.000 | 0 |
| 1/8 | 3.175 | 2 |
| 2/8 | 6.350 | 5 |
| 3/8 | 9.525 | 12 |
| 4/8 | 12.700 | 16 |
| 5/8 | 15.875 | 20 |
| 6/8 | 19.050 | 24 |

Table 6.16: Pothole deduct table (GDOT, 2007)

| Number of potholes, patches and local base failures | Deduct Value |
|---:|---:|
| 1-2 | 2 |
| 3-6 | 5 |
| 7-10 | 10 |
| 11-15 | 17 |
| >15 | 25 |

Potholes are generally individually counted and summed up for each 1-mile segment (GDOT, 2007; Miller and Bellinger, 2014). A deduct value is then assigned based on the total number of potholes per mile (table 6.16). In the proposed model, the output provides the probability of the presence of a pothole, which can be thresholded to determine whether or not the 5 meter segment has a pothole. The potholes detected in a 1-mile segment can then be counted. The lowest MAE achieved was 4 potholes per mile at a threshold of 0.052 for multiple-run pothole estimation, 4 potholes per mile at a threshold of 0.051 for pothole prediction using the known variance model and 2 potholes per mile at a threshold of 0.033 for the unknown variance model. We can see from table 6.16 that a count that is off by 4 50% of the time is not ideal for practical application.

Potholes are a unique type of distress in the sense that drivers often try to actively avoid potholes. As a result, a pothole is often absent in the vehicle vibration in several runs. This has an adverse effect due to the proposed methodology for combining multiple-run data presented in this study, which gives higher confidence to consistent results and takes the weighted average of the results across several runs. Thus, the methodology proposed in

this thesis is not suitable for pothole estimation and prediction.

This thesis focused on scalar distress values. However, cracking and raveling require detailed measurement of the extent and severity level of the distress (GDOT, 2007; FDOT, 2017) to calculate their deduct value. For cracking, this thesis focused on the crack severity level. In the PACES protocol, cracking is also reported for a 100 ft (30 m) section, which represents the condition for a 1-mile segment. After smoothing with a 100 ft window, multiple-run pavement condition estimation classifies 29% of the segments correctly (classification accuracy), while pavement condition prediction using the known and unknown variance models achieve classification accuracies of 46% and 42% respectively. Raveling is generally reported for 1-mile segments (GDOT, 2007; Miller and Bellinger, 2014). After smoothing, the MAE for the raveling index was 51.27 $cm^3/m^2$ for multiple-run raveling estimation, while that for raveling prediction was 14.24 $cm^3/m^2$ for the known variance model and 31.75 $cm^3/m^2$ for the unknown variance model.

# CHAPTER 7
# CONCLUSIONS AND RECOMMENDATATIONS

## 7.1 Contributions

This thesis contributed the following:

1. A data collection and registration method to register both smartphone data collected at fixed time intervals and 3D pavement data collected at fixed distance intervals onto a common GIS model of the road network was proposed and developed. The proposed methodology was then validated using data collected in Atlanta, GA representing various pavement conditions and road classifications.

    (a) A data collection app was developed for smartphone data collection on Android and iOS devices. The data collection app was able to collect and store data from multiple smartphone sensors along with location data with acceptable power consumption, data transfer, storage and computation requirements for mobile devices, providing a proof-of-concept that pavement condition data can be practically crowdsourced using low-cost devices such as smartphones.

    (b) The proposed methodology was designed to accommodate different frequencies of smartphone sensor (accelerometer, gyroscope, magnetometer) data collection and GPS location data collection by smartphones by synchronizing the sensors and GPS by time and interpolating the location for every sensor point.

    (c) Noise in the smartphone GPS data causes erratic movements to be recorded when the vehicle is stopped. The proposed methodology overcomes this issue by using monotonically increasing linear referencing for smartphone data registration.

Qualitative and quantitative validation of the proposed methodology for smartphone and 3D pavement data registration proved the effectiveness of the proposed methodology to successfully register the two sources of data onto a common GIS model without noise.

2. A methodology for pavement condition estimation using a single run of smartphone collected data was proposed and implemented. The proposed methodology addressed two important research needs for pavement condition estimation using smartphone data:

   (a) User context has been a large obstacle for crowdsourcing of pavement condition data. Factors such as vehicle model, vehicle speed, smartphone position, etc. do not change the pavement condition, but do have a significant impact on the sensor data captured by smartphone sensors. This makes it extremely difficult to calibrate the sensors to accurately predict the pavement condition.

   In this thesis, data from various scenarios of vehicle type, pavement condition, road classification, smartphone type and location were collected for training a deep learning model for pavement condition estimation. It was shown that the trained model had significantly better performance than the popular half-car model based method for vibration-based IRI estimation in estimating the IRI for the same road segments using different vehicles, drivers and smartphone positions.

   (b) The possibility of crowdsourcing pavement condition information using smartphones has been explored numerous times, but existing literature has primarily focused on only two distress measures: IRI and potholes. In this thesis, the possibility of detecting important distresses such as cracking, rutting and raveling using smartphone data has been explored for the first time.

3. A methodology for combining single-run pavement condition estimates to provide

a combined pavement condition estimate with a confidence level was proposed and implemented. The proposed methodology addressed two important research needs:

(a) Currently, pavement condition surveys are generally carried out only once for a given time step. With crowdsourcing of pavement condition data using smartphones, it is expected that multiple runs of smartphone data could be collected for the same road segments over a period of time. Moreover, the crowdsourced pavement data may overlap by more accurate surveys of the pavement condition. The pavement condition for the road segments can be estimated from each of these runs, but a method is needed to provide a combined pavement condition estimate with the appropriate weightage given to each source of data.

    In this thesis, a method to register and combine pavement condition estimates from multiple sources was developed. The proposed method was able to mitigate spurious pavement condition estimates from some runs and helped to lower the overall error.

(b) It was observed that the error distribution was highly skewed, i.e. most of the error was contributed by a small number of road segments with very large errors. Thus, a confidence level is needed for each pavement condition estimate which can be used to identify and reject the outlier pavement condition estimates.

    The proposed methodology provided a confidence level associated with each combined pavement condition estimate. The confidence level helped to isolate and remove outlier pavement condition estimates, decreasing the overall error.

The proposed methodology for combining pavement condition estimates is completely independent of the methodology used for estimating the single-run pavement condition. In this thesis, the method proposed in chapter 4 was used, but it can be easily replaced with another method for single-run pavement condition estimate, making this approach more flexible.

4. A methodology for pavement condition prediction using 3D pavement and smartphone data was proposed and developed.

   (a) The proposed methodology was developed for a generic distress value describing the pavement condition. Thus, this methodology can be applied to any measure of pavement distress which can possibly be captured using crowdsourced smartphone data. IRI was taken as an example of a distress value and the performance of the proposed method was tested in various road conditions and classifications. The proposed methodology obtained a significant improvement in predicting the pavement condition by leveraging information from previous time steps about the pavement condition.

   (b) Two separate models for pavement condition prediction were developed and tested. One model assumed that the variance of the distress values reported from single-run pavement condition estimates was a fixed quantity. The second model accommodates the case where this assumption is not true, and the variance is inverse gamma distributed.

## 7.2 Findings

This research made the following findings:

1. IRI estimation using 3D pavement data was validated by comparing the IRI estimated by 3D pavement data collected by the GTSV and data collected by AASHTO R56 (AASHTO, 2014) certified GDOT profilers. A mere difference of -0.58 $in/mi$ (0.92% difference) for the left wheel path and 2.57 $in/mi$ (4.2% difference) for the right wheel path IRI estimates was found.

2. Deep learning based models for estimation of various pavement distresses were trained. The performance of these models on the test set is given in table 4.2. The performance of these models in practical applications is discussed in section 6.7. Although

the desirable margin of error is different for different transportation agencies, the obtained values indicate that smartphone data can be successfully used to provide a useful estimate of pavement IRI and rutting.

3. The IRI estimated using the proposed deep learning model for single-run pavement condition estimation was found to have a much lower RMSE difference when calculated using two different vehicles as compared to that obtained from using the popular pseudo-IRI model for IRI estimation using vertical acceleration, indicating that the proposed methodology is much more robust against user context (section 4.2).

4. A desired negative correlation was found between the confidence level of IRI estimates using smartphone data and the RMSE of the estimations (section 5.3). The usefulness of the confidence level to separate and reject bad estimates was plotted in figure 5.10.

5. The use of prior information to improve pavement condition estimation was found to significantly improve estimation of IRI (table 5.1). A 5.3% reduction in overall RMSE was observed.

6. If a weightage of 1 is assigned to 3D pavement data, then an optimal weightage $\epsilon$ of 0.0450 was determined for smartphone data.

### 7.3 Limitations and Recommendations for Future Work

The following limitations and recommendations for future research were observed:

1. Although the data collection route used to collect the training data for the proposed methodology for single run pavement condition estimation attempted to cover diverse pavement conditions and road classifications, there still exist diverse scenarios which could not be covered within the scope of this study. Some examples include

sensor data from large vehicles such as buses and trucks, sensor data on continually reinforced concrete pavements (CRCP) and sensor data from roads on rolling or mountainous vertical grades. The methodology presented in this study can be used to train on additional data to create better models for single-run pavement condition estimation. The proposed methodology also has a modular network architecture, which allows the complexity of the model to be increased if required.

2. A detailed calibration procedure and eligibility criteria with minimum requirements should be established for the use of data from additional scenarios such as new vehicle types, pavement types and devices. For example, if we want to retrain the models to use trucks as well, a test section similar to the data collection route in this study needs to be established with known ground truth pavement condition which can be used to collect truck-mounted smartphone data for retraining the models. For an example of the eligibility criteria, motorcycles cannot be used for IRI estimation as IRI is defined underneath the left and right wheelpath in lanes, which motorcycles will not necessary follow.

3. The methodology for single-run pavement condition estimation was kept generalized to allow it to be extended to other distresses. However, domain knowledge of the vibration responses of specific pavement distresses can be used to precompute and provide useful features and/or postprocess the results of single-run pavement condition estimation to improve performance for those distresses.

    (a) For example, the pseudo-IRI value can be computed and provided as an additional feature dimension when estimating IRI.

    (b) Another example is the use of lateral acceleration of vehicles (possibly indicating movement to avoid potholes) to detect potholes.

    (c) Domain knowledge can also be used to postprocess the pavement condition estimates, such as blacklisting of false positive locations for pothole detection

(Eriksson, Girod, Hull, Newton, Madden, and Balakrishnan, 2008; Chen, Lu, Tan, and Wu, 2013).

(d) In this thesis, the pavement condition was discretized to 5 meter intervals. The continuity of vibration responses inherent to some pavement distresses was not considered. Analysis of the distress values estimated in adjacent 5 meter segments could potentially reveal new vibration response characteristics of these distresses, which can be used to postprocess the single-run pavement condition estimates to improve performance.

4. Estimation of five distresses: IRI, potholes, cracking, rutting and raveling were explored in this thesis. There are more measures of pavement condition, such as faulting, spalling and skid resistance, which can affect vehicle vibration and thus, be picked up by smartphone sensors. The single-run pavement condition estimation methodology presented in this thesis can be extended to train models for detecting these additional distresses as well.

5. In this thesis, data registration and processing was carried out after transferring the data off the device. This may raise privacy concerns as the data includes location data from the smartphone. There are two possible solutions which can be explored in future research:

   (a) Anonymization and obfuscation of data before transferring it out of the device can protect users' privacy, while still allowing parts of the data to be used for pavement condition estimation.

   (b) With the increasing processing capabilities of smartphones, the model can be stored on the devices themselves and the single-run pavement condition estimation can be carried out on the device itself. Thus, the raw location data does not need to be transferred.

6. In this thesis, video data was not considered as the power consumption, data transfer time and costs would make it infeasible to crowdsource. However, the capabilities of smartphones have grown rapidly in the past decade. With the advent of 5G, the data transfer time issue could be potentially solved. Thus, in the future, the use of computer vision to detect additional road infrastructure condition information, such as traffic sign condition, guardrail condition, etc. can be explored.

7. The distribution of typical values for a given distress measure can be used to filter out outliers in multiple-run data for a road segment before multiple-run pavement condition estimation and confidence level calculation, which was not explored in this thesis. For example, if a large number of runs are collected for a road segment, single-run pavement condition estimates below the 25th percentile and above the 75th percentile values can be rejected before calculating the combined pavement condition estimate and confidence level.

8. The proposed methodology in this thesis can also be applied to low-cost sensors on probe vehicles and connected vehicles with built-in sensors as an alternative to smartphones. Embedded sensors attached close to the vehicle axles may provide better information about the pavement condition before the vibration response is dampened by the vehicle suspension.

## REFERENCES


[1] FHWA, "Chart VMT-422 - Highway Statistics 2017," Federal Highway Administration, Tech. Rep. VMT-421C, 2017.

[2] N. Musick, *Public spending on transportation and water infrastructure, 1956 to 2017*, https://www.cbo.gov/system/files/2018-10/54539-Infrastructure.pdf, Accessed: June 2, 2019, 2018.

[3] ASCE, *Infrastructure report card: Roads*, https://www.infrastructurereportcard.org/wp-content/uploads/2017/01/Roads-Final.pdf, Accessed: November 8, 2018, 2017.

[4] The White House, *Legislative outline for rebuilding infrastructure in america*, https://www.whitehouse.gov/wp-content/uploads/2018/02/INFRASTRUCTURE-211.pdf, Accessed: June 2, 2019, 2018.

[5] GDOT, "Pavement condition evaluation system," Georgia Department of Transportation, Tech. Rep., 2007.

[6] G. Rada, R. Bhandari, G. Elkins, and W. Bellinger, "Assessment of long-term pavement performance program manual distress data variability: Bias and precision," *Transportation Research Record: Journal of the Transportation Research Board*, no. 1592, pp. 151–168, 1997.

[7] Y. Tsai, Z. Wang, *et al.*, "A remote sensing and gis-enabled asset management system (rs-gams).," 2013.

[8] FHWA, *Tpf-5(299): Pavement surface condition data and analysis*, https://collaboration.fhwa.dot.gov/dot/fhwa/tpf5299/default.aspx, Accessed: January 1, 2019, 2017.

[9] J. Dawkins, R. Bishop, B. Powell, and D. Bevly, "Investigation of pavement maintenance applications of intellidrive sm (final report): Implementation and deployment factors for vehicle probe-based pavement maintenance (pbpm)," *Auburn University*, 2011.

[10] P. M. Sauerwein, B. L. Smith, *et al.*, "Investigation of the implementation of a probe-vehicle based pavement roughness estimation system.," Virginia Transportation Research Council, Tech. Rep., 2011.

[11] G. Flintsch, S. Valeri, S. Katicha, E. de Leon Izeppi, and A. Medina-Flintsch, "Probe vehicles used to measure road ride quality: Pilot demonstration," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2304, pp. 158–165, 2012.

[12] K. Zang, J. Shen, H. Huang, M. Wan, and J. Shi, "Assessing and mapping of road surface roughness based on gps and accelerometer sensors on bicycle-mounted smartphones," *Sensors*, vol. 18, no. 3, p. 914, 2018.

[13] J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden, and H. Balakrishnan, "The pothole patrol: Using a mobile sensor network for road surface monitoring," in *Proceedings of the 6th international conference on Mobile systems, applications, and services*, ACM, 2008, pp. 29–39.

[14] K. Chen, M. Lu, G. Tan, and J. Wu, "Crsm: Crowdsourcing based road surface monitoring," in *2013 IEEE 10th International Conference on High Performance Computing and Communications & 2013 IEEE International Conference on Embedded and Ubiquitous Computing*, IEEE, 2013, pp. 2151–2158.

[15] M. W. Sayers and S. M. Karamihas, "Interpretation of road roughness profile data," Tech. Rep. FHWA/RD-96/101, 1996.

[16] H Zeng, "Improving pavement management and assessment through connected vehicle technology and highway safety analysis," PhD thesis, University of Virginia, 2014.

[17] H. Zeng, H. Park, B. L. Smith, and E. Parkany, "Feasibility assessment of a smartphone-based application to estimate road roughness," *KSCE Journal of Civil Engineering*, vol. 22, no. 8, pp. 3120–3129, 2018.

[18] Y. Tsai and Z. Wang, "A remote sensing and gis-enabled asset management system (rs-gams): Phase 2.," United States. Dept. of Transportation. Office of the Assistant Secretary, Tech. Rep., 2014.

[19] A. Chatterjee, "A fast and accurate automated pavement crack detection algorithm," PhD thesis, Georgia Institute of Technology, 2017.

[20] Y.-C. J. Tsai and F. Li, "Critical assessment of detecting asphalt pavement cracks under different lighting and low intensity contrast conditions using emerging 3d laser technology," *Journal of Transportation Engineering*, vol. 138, no. 5, pp. 649–656, 2012.

[21] C. Jiang and Y. J. Tsai, "Enhanced crack segmentation algorithm using 3d pavement data," *Journal of Computing in Civil Engineering*, vol. 30, no. 3, p. 04 015 050, 2016.

[22] J. Y.-C. Tsai, Z.-H. Wang, and F. Li, "Assessment of rut depth measurement accuracy of point-based rut bar systems using emerging 3d line laser imaging technology," *Journal of Marine Science and Technology*, vol. 23, no. 3, pp. 322–330, 2015.

[23] J. Y.-C. Tsai, F. Li, and Y.-C. Wu, "A new rutting measurement method using emerging 3d line-laser-imaging system," *International Journal of Pavement Research and Technology*, vol. 6, no. 5, pp. 667–672, 2013.

[24] Y. J. Tsai and Z. Wang, "Development of an asphalt pavement raveling detection algorithm using emerging 3d laser technology and macrotexture analysis," Tech. Rep. 20-30/IDEA 163, 2015.

[25] Y.-C. Tsai and A. Chatterjee, "Pothole detection and classification using 3d technology and watershed method," *Journal of Computing in Civil Engineering*, vol. 32, no. 2, p. 04 017 078, 2017.

[26] Y. J. Tsai, Y. Wu, and C. Ai, "Feasibility study of measuring concrete joint faulting using 3d continuous pavement profile data 2," in *Proceedings of the Transportation Research Board 90th Annual Meeting, Washington, DC, USA*, 2011, pp. 23–27.

[27] Y. Tsai, Y. Wu, and Z. Lewis, "Full-lane coverage micromilling pavement-surface quality control using emerging 3d line laser imaging technology," *Journal of Transportation Engineering*, vol. 140, no. 2, p. 04 013 006, 2013.

[28] K. H. McGhee, *Automated pavement distress collection techniques*. Transportation Research Board, 2004, vol. 334.

[29] B. Xu, S Ranji Ranjithan, and Y Richard Kim, "New relationships between falling weight deflectometer deflections and asphalt pavement layer condition indicators," *Transportation Research Record: Journal of the Transportation Research Board*, no. 1806, pp. 48–56, 2002.

[30] O. Elbagalati, M. Elseifi, K. Gaspard, and Z. Zhang, "Development of the pavement structural health index based on falling weight deflectometer testing," *International Journal of Pavement Engineering*, vol. 19, no. 1, pp. 1–8, 2018.

[31] H. S. Abd El-Raof, R. T. Abd El-Hakim, S. M. El-Badawy, and H. Afify, "Structural number prediction for flexible pavements based on falling weight deflectometer data," Tech. Rep., 2018.

[32] W. W.-L. Lai, X. Dérobert, and P. Annan, "A review of ground penetrating radar application in civil engineering: A 30-year journey from locating and testing to imaging and diagnosis," *NDT & E International*, vol. 96, pp. 58–78, 2018.

[33] M. Nasimifar, S. Thyagarajan, R. V. Siddharthan, and N. Sivaneswaran, "Robust deflection indices from traffic-speed deflectometer measurements to predict critical pavement responses for network-level pavement management system application," *Journal of Transportation Engineering*, vol. 142, no. 3, p. 04 016 004, 2016.

[34] A. Zofka and J. Sudyka, "Traffic speed deflectometer (tsd) measurements for pavement evaluation," in *International Symposium Non-destructive Testing in Civil Engineering (NDTCE)*, 2015.

[35] K. C. Wang, "Designs and implementations of automated systems for pavement surface distress survey," *Journal of Infrastructure Systems*, vol. 6, no. 1, pp. 24–32, 2000.

[36] K. C. Wang, "Challenges and feasibility for comprehensive automated survey of pavement conditions," in *Applications of Advanced Technologies in Transportation Engineering (2004)*, 2004, pp. 531–536.

[37] A. Zhang, K. C. Wang, B. Li, E. Yang, X. Dai, Y. Peng, Y. Fei, Y. Liu, J. Q. Li, and C. Chen, "Automated pixel-level pavement crack detection on 3d asphalt surfaces using a deep-learning network," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 10, pp. 805–819, 2017.

[38] C. Koch and I. Brilakis, "Pothole detection in asphalt pavement images," *Advanced Engineering Informatics*, vol. 25, no. 3, pp. 507–515, 2011.

[39] J Laurent, D Lefebvre, and E Samson, "Development of a new 3d transverse laser profiling system for the automatic measurement of road cracks," in *Symposium on Pavement Surface Characteristics, 6th, 2008, Portoroz, Slovenia*, 2008.

[40] S. Haider, G. Baladi, K. Chatti, and C. Dean, "Effect of frequency of pavement condition data collection on performance prediction," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2153, pp. 67–80, 2010.

[41] Federal Aviation Administration, *Airport obligations: Pavement maintenance*, `https://www.faa.gov/airports/central/airport_compliance/pavement_maintenance/`, Accessed: January 1, 2019, 2017.

[42] W. Aleadelat, K. Ksaibati, C. H. Wright, and P. Saha, "Evaluation of pavement roughness using an android-based smartphone," *Journal of Transportation Engineering, Part B: Pavements*, vol. 144, no. 3, p. 04 018 033, 2018.

[43] S. Islam, W. G. Buttlar, R. G. Aldunate, and W. R. Vavrik, "Measurement of pavement roughness using android-based smartphone application," *Transportation Research Record*, vol. 2457, no. 1, pp. 30–38, 2014.

[44] Y. Yang and K. Chang, "Extracting the bridge frequencies indirectly from a passing vehicle: Parametric study," *Engineering Structures*, vol. 31, no. 10, pp. 2448–2459, 2009.

[45] Y.-B. Yang, C. Lin, and J. Yau, "Extracting bridge frequencies from the dynamic response of a passing vehicle," *Journal of Sound and Vibration*, vol. 272, no. 3-5, pp. 471–493, 2004.

[46] W. Aleadelat and K. Ksaibati, "Estimation of pavement serviceability index through android-based smartphone application for local roads," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2639, pp. 129–135, 2017.

[47] P. Mohan, V. N. Padmanabhan, and R. Ramjee, "Nericell: Rich monitoring of road and traffic conditions using mobile smartphones," in *Proceedings of the 6th ACM conference on Embedded network sensor systems*, ACM, 2008, pp. 323–336.

[48] R. Bhoraskar, N. Vankadhara, B. Raman, and P. Kulkarni, "Wolverine: Traffic and road condition estimation using smartphone sensors," in *2012 Fourth International Conference on Communication Systems and Networks (COMSNETS 2012)*, IEEE, 2012, pp. 1–6.

[49] A. Thiagarajan, L. Ravindranath, K. LaCurts, S. Madden, H. Balakrishnan, S. Toledo, and J. Eriksson, "Vtrack: Accurate, energy-aware road traffic delay estimation using mobile phones," in *Proceedings of the 7th ACM conference on embedded networked sensor systems*, ACM, 2009, pp. 85–98.

[50] A. Vittorio, V. Rosolino, I. Teresa, C. M. Vittoria, P. G. Vincenzo, *et al.*, "Automated sensing system for monitoring of road surface quality by mobile devices," *Procedia-Social and Behavioral Sciences*, vol. 111, pp. 242–251, 2014.

[51] A. Ghose, P. Biswas, C. Bhaumik, M. Sharma, A. Pal, and A. Jha, "Road condition monitoring and alert application: Using in-vehicle smartphone as internet-connected sensor," in *2012 IEEE International Conference on Pervasive Computing and Communications Workshops*, IEEE, 2012, pp. 489–491.

[52] V. Astarita, M. V. Caruso, G. Danieli, D. C. Festa, V. P. Giofrè, T. Iuele, and R. Vaiana, "A mobile application for road surface quality control: Uniqualroad," *Procedia-Social and Behavioral Sciences*, vol. 54, pp. 1135–1144, 2012.

[53] A. Massahi, H. Ali, F. Koohifar, M. Baqersad, and M. Mohammadafzali, "Implementing a smartphone application for assessing the raveling performance of asphalt pavements," *International Journal of Advances in Scientific Research and Engineering*, vol. 4, 2018.

[54] ——, "Investigation of pavement raveling performance using smartphone," *International Journal of Pavement Research and Technology*, vol. 11, no. 6, pp. 553–563, 2018.

[55] J. Takahashi, Y. Kobana, N. Isoyama, Y. Tobe, and G. Lopez, "Ykob: Participatory sensing based road condition monitoring using smartphones worn by cyclist," *IEEJ transactions on electronics, information and systems*, vol. 137, no. 4, pp. 658–666, 2017.

[56] M. Ndoye, A. M. Barker, J. V. Krogmeier, and D. M. Bullock, "A recursive multi-scale correlation-averaging algorithm for an automated distributed road-condition-monitoring system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 3, pp. 795–808, 2011.

[57] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.

[58] Q.-J. Kong, Q. Zhao, C. Wei, and Y. Liu, "Efficient traffic state estimation for large-scale urban road networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 1, pp. 398–407, 2013.

[59] R. Herring, A. Hofleitner, P. Abbeel, and A. Bayen, "Estimating arterial traffic conditions using sparse probe data," in *13th International IEEE Conference on Intelligent Transportation Systems*, IEEE, 2010, pp. 929–936.

[60] A. Hadachi, C. Lecomte, S. Mousset, and A. Bensrhair, "An application of the sequential monte carlo to increase the accuracy of travel time estimation in urban areas," in *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, IEEE, 2011, pp. 157–162.

[61] K. Aihara, H. Imura, A. Takasu, Y. Tanaka, and J. Adachi, "Crowdsourced mobile sensing for smarter city life," in *2014 IEEE 7th International Conference on Service-Oriented Computing and Applications*, IEEE, 2014, pp. 334–337.

[62] L. Tang, X. Yang, Z. Dong, and Q. Li, "Clric: Collecting lane-based road information via crowdsourcing," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 9, pp. 2552–2562, 2016.

[63] M. Byrne, R. Isola, and T. Parry, "A generalised approach to outlier identification in pavement condition data," *International Journal of Pavement Engineering*, vol. 14, no. 1, pp. 60–70, 2013.

[64] *Roadbotics*, `http://www.roadbotics.com`, Accessed: June 11, 2019, 2016.

[65] *Roadroid*, `http://www.roadroid.com`, Accessed: June 11, 2019.

[66] M Scholotjes, A Visser, and C Bennett, "Evaluation of a smartphone roughness meter," in *33rd Southern African Transport Conference (SATC 2014)*, Citeseer, 2014.

[67] V. Douangphachanh and H. Oneyama, "A study on the use of smartphones for road roughness condition estimation," *Journal of the Eastern Asia Society for Transportation Studies*, vol. 10, pp. 1551–1564, 2013.

[68] ——, "Estimation of road roughness condition from smartphones under realistic settings," in *2013 13th international conference on ITS telecommunications (ITST)*, IEEE, 2013, pp. 433–439.

[69] ——, "Using smartphones to estimate road pavement condition," 2013.

[70] T. Nagayama, A. Miyajima, S. Kimura, Y. Shimada, and Y. Fujino, "Road condition evaluation using the vibration response of ordinary vehicles and synchronously recorded movies," in *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2013*, International Society for Optics and Photonics, vol. 8692, 2013, 86923A.

[71] FDOT, "2017 flexible pavement condition survey handbook," Florida Department of Transportation, Tech. Rep., 2017.

[72] F. N. Fritsch and R. E. Carlson, "Monotone piecewise cubic interpolation," *SIAM Journal on Numerical Analysis*, vol. 17, no. 2, pp. 238–246, 1980.

[73] R. Robinson, "Slippery road detection and evaluation," 2012.

[74] L. Sun, "Simulation of pavement roughness and iri based on power spectral density," *Mathematics and computers in simulation*, vol. 61, no. 2, pp. 77–88, 2003.

[75] J. Laurent, J. F. Hébert, D. Lefebvre, and Y. Savard, "Using 3d laser profiling sensors for the automated measurement of road surface conditions," in *7th RILEM international conference on cracking in pavements*, Springer, 2012, pp. 157–167.

[76] J. S. Miller, W. Y. Bellinger, *et al.*, "Distress identification manual for the long-term pavement performance program," United States. Federal Highway Administration. Office of Infrastructure, Tech. Rep., 2014.

[77] J. M. Williams, "Deep learning and transfer learning in the classification of EEG signals," PhD thesis, 2017.

[78] R. V. Lenth, "Some practical guidelines for effective sample size determination," *The American Statistician*, vol. 55, no. 3, pp. 187–193, 2001.

[79] AASHTO, "Standard practice for certification of inertial profiling systems," American Association of State Highway Transportation Officials, Tech. Rep. R56, 2014.

[80] FHWA, "Measuring and specifying pavement smoothness," Federal Highway Administration, Tech. Rep. FHWA-HIF-16-032, 2016.